Semantic or Affective Primacy?

Perceptual Latencies of Object Recognition and Affect Measured With Temporal Judgments and Speeded Reaction Time Tasks

Inauguraldissertation

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von

Dipl.-Psych. Philipp Franikowski

Greifswald, 11. Oktober 2022

Dekan: Prof. Dr. Gerald Kerth

1. Gutachter: Prof. Dr. Rainer Reisenzein

2. Gutachter: Prof. Dr. Dirk Wentura

3. Gutachter: Prof. Dr. Maurizio Codispoti

Tag der Promotion: 02.03.2023

Abstract

A long-standing controversy in emotion research concerns the question whether stimuli must be conceptually interpreted, or semantically categorized, to evoke emotional reactions. According to the semantic primacy hypothesis, the answer to this question is positive; whereas according to the affective primacy hypothesis, it is negative: Emotions can also be, and perhaps often are, elicited by preconceptual stimulus representations, such as particular shapes or color patterns.

In the present dissertation project, the semantic primacy hypothesis was tested in eight experiments using different latency judgment paradigms in which the perceptual latencies of object recognition and affect onset were measured and compared. The chronometric measurement methods comprised temporal judgments (temporal order judgments and simultaneity judgments: Publication A, Experiments 1–4; the rotating spot / rotating clock hand method: Publication B, Experiments 1–2) and speeded reaction time measurements (Publication C, Experiments 1–2). To elicit affective responses, pictures of pleasant (e.g., cats, children) and unpleasant objects (e.g., spiders, moldy food) from everyday life were presented.

According to the semantic primacy hypothesis, object recognition is a necessary partial cause of affect. This implies the following three predictions that were tested in the studies: (1) Because causes must precede their effects, the time of the onset of object recognition must precede the time of the onset of affect. (2) The longer it takes a person to recognize an object, the longer it should also take them, other factors constant, to experience affect; therefore, the latencies of the two mental events should be positively correlated across individuals. (3) An experimental manipulation that delays the onset of object recognition (in this case a moderate blurring of the pictures) should also delay the onset of affect, and the effect of the manipulation on affect latency should be mediated by the delay in object recognition.

In agreement with Prediction 1, regardless of the chronometric method used, the latency of object recognition consistently proved to be shorter than the latency of affect onset. According to the meta-analytically integrated latency differences estimated in the temporal judgment experiments, affect followed object recognition with a delay of 117 ms. This result was obtained for both pleasant and unpleasant stimuli and was independent of task order. Supporting Prediction 2, the latencies for object recognition and affect onset were positively correlated across participants (meta-analytic $r = .50$). Supporting Prediction 3, delaying object recognition by blurring the affective pictures was found to also delay the onset of affect and the effect of blurring on the latency of affect was found to be partly mediated by delayed object recognition.

Two additional predictions tested and confirmed in Experiment C2 were: (4) False-coloring the affective pictures delays the onset of affect but not object recognition, and this effect is mediated by reduced affect intensity. (5) Judgments of the valence of the stimuli (i.e., whether the imaged object is pleasant or unpleasant) take more time than reports of object recognition, but less time than affect onset reports, for which valence judgments have often been used as a substitute in previous studies.

Taken together, the results of the eight experiments provided consistent support for semantic primacy in the generation of pleasant and unpleasant feelings evoked by affective pictures: Object recognition can be considered a necessary partial cause of affect in the reported experiments. The results are compared to previous

findings, possible reasons for deviant response patterns found in a small minority of the participants are considered, and several implications of the findings for emotion research are derived. Possible adaptations of the chronometric approach to investigate other questions of emotion research are suggested. Finally, limitations of the dissertation project are pointed out and possible ways to address these in future research are proposed.

Zusammenfassung

Eine seit langem bestehende Kontroverse in der Emotionsforschung betrifft die Frage, ob Stimuli konzeptuell interpretiert oder semantisch kategorisiert werden müssen, um emotionale Reaktionen hervorzurufen. Nach der Hypothese des semantischen Primats (Semantic-Primacy-Hypothese) ist diese Frage positiv zu beantworten; nach der Hypothese des affektiven Primats (Affective-Primacy-Hypothese) hingegen ist sie negativ zu beantworten: Emotionen können auch durch präkonzeptuelle Stimulusrepräsentationen, wie z. B. bestimmte Formen oder Farbmuster, ausgelöst werden – und werden es vermutlich sogar oft.

Die Semantic-Primacy-Hypothese wurde im Rahmen dieses Dissertationsprojekts in acht Experimenten anhand von verschiedenen Latenzmessungsparadigmen getestet, in denen die Wahrnehmungslatenzen von Objekterkennung einerseits und von Affektauslösung andererseits gemessen und miteinander verglichen wurden. Die chronometrischen Messmethoden umfassten temporale Urteile (Temporal Order Judgments und Simultaneity Judgments: Publikation A, Experimente 1–4; die Rotating-Spot- bzw. Uhrzeigermethode: Publikation B, Experimente 1–2) sowie Reaktionszeitmessungen (Publikation C, Experimente 1–2). Um affektive Reaktionen hervorzurufen, wurden Bilder von angenehmen (z. B. Katzen, Kinder) und unangenehmen Objekten (z. B. Spinnen, verdorbene Lebensmittel) aus dem alltäglichen Leben präsentiert.

Nach der Semantic-Primacy-Hypothese ist die Objekterkennung eine notwendige Teilursache des Affekts. Dies impliziert die folgenden drei Vorhersagen, die in den Studien getestet wurden: (1) Da im Sinne der Kausalität Ursachen ihren Wirkungen vorausgehen, muss der Zeitpunkt des Einsetzens der Objekterkennung vor dem Zeitpunkt des Affekts liegen. (2) Je länger eine Person braucht, um ein Objekt zu erkennen, desto länger sollte sie, unter Konstanz aller anderen Faktoren, auch brauchen, um einen Affekt zu erleben; daher sollten die Latenzen der beiden mentalen Ereignisse über alle Personen hinweg positiv korreliert sein. (3) Eine experimentelle Manipulation, die das Einsetzen der Objekterkennung verzögert (wir wendeten einen moderaten Unschärfefilter auf die affektiven Bilder an), sollte auch das Einsetzen des Affekts verzögern und der Effekt der Manipulation auf den Affekt sollte wiederum durch die Verzögerung der Objekterkennung mediiert sein.

In Übereinstimmung mit Vorhersage 1 war die Latenzzeit der Objekterkennung unabhängig von der verwendeten chronometrischen Methode durchweg kürzer als die Latenzzeit des Affekts. Den meta-analytisch integrierten Latenzdifferenzen entsprechend, die in den Experimenten mit temporalen Urteilen geschätzt wurden, folgte der Affekt 117 ms nach der Objektrekognition. Dieses Ergebnis zeigte sich sowohl für angenehme als auch für unangenehme Stimuli und war unabhängig von der Aufgabenreihenfolge. Wie durch Vorhersage 2 nahegelegt, waren die Latenzen von Objekterkennung und Affekt über alle Versuchspersonen hinweg positiv korreliert (meta-analytisches $r = .50$). Vorhersage 3 stützend wurde festgestellt, dass die Verzögerung der Objekterkennung durch Anwendung des Unschärfefilters auf die affektiven Bilder auch das Auftreten des Affekts verzögert und dass der Effekt des Unschärfefilters auf die Latenz des Affekts teilweise durch die verzögerte Objekterkennung mediiert ist.

Zwei weitere Vorhersagen wurden in Experiment C2 getestet und bestätigt: (4) Eine Falscheinfärbung der affektiven Bilder verzögert das Einsetzen des Affekts,

aber nicht die Objekterkennung, und dieser Effekt wird durch eine verringerte Affektintensität mediiert. (5) Urteile über die Valenz der Stimuli (d. h., ob das abgebildete Objekt angenehm oder unangenehm ist) benötigen mehr Zeit als Berichte über die Objekterkennung, aber weniger Zeit als Berichte über das Einsetzen des Affekts. Für Berichte des Einsetzens des Affekts wurden Valenzurteile in früheren Studien oft stellvertretend herangezogen.

Zusammengenommen unterstützen die Ergebnisse der acht Experimente konsistent die Semantic-Primacy-Hypothese der Entstehung angenehmer und unangenehmer Gefühle durch affektive Bilder: Objektrekognition kann als notwendige Teilursache des Affekts in den berichteten Experimenten betrachtet werden. Die Ergebnisse werden mit früheren Befunden verglichen, mögliche Gründe für abweichende Ergebnisse bei einer kleinen Gruppe der Versuchspersonen erörtert und verschiedene Implikationen der Ergebnisse für die Emotionsforschung abgeleitet. Mögliche Anpassungen des chronometrischen Ansatzes zur Untersuchung anderer Fragestellungen in der Emotionsforschung werden vorgeschlagen. Abschließend wird auf die Grenzen des Dissertationsprojekts eingegangen und es werden Möglichkeiten vorgeschlagen, wie diese in zukünftiger Forschung adressiert werden können.

Contents

Chapter 1

Introduction

At the sight of a cute little puppy or a kitten, most people experience pleasure. At the sight of a person vomiting or a large spider, on the other hand, most people experience displeasure – even disgust, in these cases. But how do objects manage to evoke pleasant and unpleasant feelings? In particular, is it necessary to identify the objects that trigger the feeling? That is, do we need to recognize that an object in front of us is a puppy to feel pleasure, or a spider to feel disgust?

Cognitively oriented emotion theorists (e.g., Lazarus, 1982; Ortony et al., 1988; Scherer, 2009; Storbeck et al., 2006) answer this question positively. These theorists endorse what has been called the *semantic primacy hypothesis* (e.g., Storbeck et al., 2006), according to which object-evoked emotions, like all emotions, presuppose a substantial amount of information processing; in particular, they require the recognition (semantic classification) of the eliciting object and its positive or negative evaluation. Recognizing an object as being a puppy means that the sensory information received from the object activates the concept PUPPY or BABY DOG (or a sufficiently related concept) in semantic memory and that this concept is then applied to the perceived object. This, however, is not enough for experiencing pleasure. In addition, to evoke pleasure, the perceived puppy must also be evaluated as being positive in some way (e.g., as cute).

Plausible as it may seem, the semantic primacy hypothesis has been disputed by several prominent emotion theorists (e.g., Cannon, 1927; James, 1890; LeDoux, 1998; Öhman & Mineka, 2001; Watson, 1919; Zajonc, 1980) who proposed instead the *affective primacy hypothesis*. These "noncognitive" emotion theorists argue that affect (pleasant or unpleasant feeling; e.g., Barrett & Bliss–Moreau, 2009) can be, and often is, triggered without much cognitive processing; in particular, it can be triggered in an automatic, reflex-like manner by the mere perception of the eliciting object without, or prior to, its semantic classification or recognition. In the example, this means that it is not necessary for feeling pleasure to be aware that the object depicted in a photograph is a puppy. In our view, the most plausible explication of the affective primacy hypothesis (Zajonc, 1980) is that affective reactions are evoked by *preconceptual* or *nonconceptual* representations of the eliciting object, such as image-like representations of the object's shape or color patterns (in the case of the puppy, e.g., its fluffiness) – at least in some cases, if not in many, or even all.

These opposing views on the elicitation of pleasant and unpleasant affect (and emotions more generally) have led to a debate between "noncognitivists" and "cognitivists" in emotion research that is still ongoing.[1] In this debate, both sides have, in addition to theoretical arguments for their position, referred to a variety of empirical data, including experimental tests of the semantic versus affective primacy hypothesis. Most of these experimental studies have used the so-called subliminal perception paradigm (in one version also referred to as "subliminal affective priming"; Murphy & Zajonc, 1993). Whereas noncognitive emotion theorists have taken the results of these studies to support the affective primacy hypothesis, their cognitivist critics have pointed out the theoretical problems and methodological shortcomings of these studies. Indeed, studies that controlled for these methodological problems obtained results that are better in line with semantic primacy (e.g., Codispoti et al., 2006; Lähteenmäki et al., 2015; Peira et al., 2012; Pessoa et al., 2005).

This conclusion agrees with that suggested by a second set of studies that used a chronometric approach to test the semantic primacy hypothesis (e.g., Nakashima, 1909a, 1909b; Nummenmaa et al., 2010). At the time when this dissertation project was started, there were only a handful of studies that used the chronometric paradigm, and with the exception of a recent set of studies by Nummenmaa et al. (2010), they were over a hundred years old (Nakashima, 1909a, 1909b). In fact, the studies by Nakashima (1909a, 1909b) were the first empirical tests ever of the semantic primacy hypothesis. Although the results of these studies were in line with the semantic primacy hypothesis, they also had a number of methodological shortcomings. In the present dissertation project, this line of research was taken up again using four different, modern methods of latency measurement and associated statistical analysis methods and much larger sets of stimuli (here, affective pictures) and participants to see whether Nakashima's findings would hold up. In addition, in two of the experiments (Experiments B2 and C2), the test of the semantic primacy hypothesis in the latency measurement paradigm was, for the first time, extended to an experimental test. In this test, object recognition was experimentally delayed (by a moderate blurring of the affective pictures) to test the effects of this manipulation on the latency of affect onset. If the semantic primacy hypothesis is correct, this manipulation should also delay the onset of affect, and furthermore, the latency of affect should be partly mediated by the latency of object recognition. Confirmation of these hypotheses would provide strong, and potentially decisive, additional support for semantic primacy.

The main part of the dissertation project consists of eight experiments distributed across three publications that tested the semantic primacy hypothesis for affect-evoking pictures. The publications are here abbreviated A, B, and C. Each publication used a different chronometric method for testing the semantic primacy hypothesis: The studies in Publication A used temporal order and simultaneity judgments, two frequently used temporal judgment methods; those in Publication B used the rotating spot method as yet another, different temporal judgment method; and Publication C reports results obtained using speeded reaction time tasks. Together, these methods comprise the main types of latency measurements of subjective events (ignoring event-related potential [ERP] measurements that index object recognition and affect only indirectly).

[1] This debate and its implications have been widely discussed not only in emotion psychology but also in other fields of psychology, as well as in philosophy (e.g., Scarantino, 2010), history (e.g., Leys, 2017) and affective computing (e.g., Marsella et al., 2010).

Chapter 2

Theoretical and Empirical Background

2.1 The Temporal Order of Object Recognition and Affect

The semantic primacy hypothesis has been advocated by cognitive emotion theorists (e.g., Lazarus, 1982; Ortony et al., 1988; Scherer, 2009; Storbeck et al., 2006). It states that emotional reactions evoked by objects (e.g., as in the introductory example, a puppy or a spider) necessarily require the recognition of the object, i.e., the recognition that the object is of a particular conceptual kind (e.g., as in the introductory example, a puppy, or at least an animal similar to a baby dog). To recognize the eliciting object thus means to subsume the object under a concept, or to classify it as an instance of a semantic category. The necessity in "necessarily require" is to be understood as causal necessity: According to the semantic primacy hypothesis, object recognition is a necessary partial cause of the affective response to an object (e.g., Arnold, 1960; Lazarus, 1995; Ortony et al., 1988). It is only a partial cause because in addition, the object also needs to be evaluated as being in some sense good or bad, e.g., as cute or as abhorrent (e.g., Arnold, 1960; Lazarus, 1995; Ortony et al., 1988).

Because causes must precede their effects, the semantic primacy hypothesis implies that object recognition temporally precedes affect, or as Zajonc (1980) summarized this implication: "Before I can like something, I must first have some knowledge about it" (p. 151). And although this is not always spelled out clearly by the proponents of semantic primacy, the reason for this is precisely that the conceptual representation of the object is the *cognitive basis* of the subsequent evaluation of the object as positive or negative that according to cognitive emotion theorists is the proximate cause of the emotional response (e.g., Lazarus, 1991). That is, the object is evaluated as positive or negative and consequently evokes an affective response, *because* it is classified as a member of a particular category.

Despite its (arguably) intuitive plausibility, the semantic primacy hypothesis has been vigorously opposed by many prominent emotion theorists beginning with James (1890), Watson (1919), and Cannon (1927). In recent times, main opponents of semantic primacy have been LeDoux (1998), Öhman and Mineka (2001), and Zajonc (1980). These authors have instead proposed what Zajonc (1980) termed the *affective primacy hypothesis*, that can be summarized in the statement: "I can like something before I know (or: without knowing) what it is." That is, the affective primacy hypothesis claims that (at least some) emotional reactions (specifically the pleasant and unpleasant feelings evoked by visually presented objects) can be (but perhaps, even regularly are; Zajonc, 1980) evoked by stimuli via a "noncognitive pathway." This means in particular that stimuli can evoke affect by a route that bypasses conceptual representation (object recognition). Instead, presumably, affect is evoked by a *preconceptual* or *nonconceptual* representation (Crane, 2009; Marr, 1982) of the stimulus, such as an image-like representation of the object having a particular form, texture, or color. For example, according to LeDoux (1998), snakes evoke negative affect as soon as an evolutionary pattern-recognition mechanism detects their S-shaped form.

2.2 Empirical Evidence

To empirically decide the semantic versus affective primacy debate, one needs to test whether object recognition is, or is not, a necessary partial cause of affect. If object recognition is not a necessary partial cause of affect (as the affective primacy hypothesis suggests), affective stimuli can elicit emotional reactions even (1) when they are not recognized or (2) before they are recognized (which is really a special case of 1). Therefore, if emotional reactions are found to occur (1a) without object recognition or (2a) before object recognition, the semantic primacy hypothesis is falsified. In the absence of a plausible third alternative, following Popper's (1959) logic of scientific discovery, the affective primacy hypothesis can be maintained for the time being. In contrast, if emotional reactions occur (1b) only in the presence of object recognition or (2b) only after object recognition, the affective primacy hypothesis is falsified. The semantic primacy hypothesis is supported and can be maintained for the time being.

In addition, the semantic primacy hypothesis predicts (3) that the latencies of object recognition and affect are positively correlated. The reason is this: If object recognition is a partial cause of affect, affect cannot occur before object recognition has occurred, and the more time a person needs to recognize the object, the longer it takes, other factors being constant, for the affect to occur. Therefore, if the two latencies are uncorrelated, this speaks against semantic primacy (and due to the lack of a third alternative, for affective primacy); whereas if they are positively correlated, this is additional support for semantic primacy.

Seen from the present perspective, the subliminal perception experiments reviewed below attempted to test prediction (1) of the semantic primacy hypothesis, i.e., the prediction that no affect occurs if object recognition is prevented. In contrast, the chronometric studies that are at the focus of this dissertation project tested prediction (2) that affect occurs only after the object was recognized, as well as prediction (3) that the two latencies are positively correlated.

2.2.1 Subliminal Perception Studies

The subliminal perception paradigm involves a very brief (about 30 ms) presentation of an affective picture with immediate backward masking. This procedure is intended to prevent the conscious recognition of the picture, while still allowing a preconceptual representation of the eliciting stimulus that evoked affect. Seemingly supporting the affective primacy hypothesis, subliminal stimulus presentations have been found to elicit peripheral physiological responses (e.g., Öhman & Soares, 1994), emotional facial expressions (e.g., Dimberg et al., 2000), and amygdala activity (e.g., Morris et al., 1998) similar to the reactions observed when the stimuli are presented above the perceptual threshold. In addition, subliminal affective stimuli were found to influence the evaluation of subsequent stimuli (affective priming; Murphy & Zajonc, 1993). If one assumes that the masking procedure did indeed prevent the recognition of the objects, only allowing the construction of a preconceptual stimulus representation, these results constitute evidence for the affective primacy hypothesis: The objects evoked affect even though they were not semantically categorized, and therefore no categorization-based evaluation could take place.

Although the proponents of the semantic primacy hypothesis have mostly accepted the empirical findings of the subliminal perception studies, they see glaring theoretical and methodological limitations of these studies. First, cognitive theorists have emphasized that object recognition – parallel to affective reactions – might have occurred at an unconscious level of processing. The subliminal perception studies do not rule this out because they measured only affective reactions but not semantic processing (Nummenmaa et al., 2010). A direct comparison of semantic and affective stimulus processing might have found that a conceptual representation of the eliciting object was in fact present and caused the affective reaction. Second, contrary to the assumptions of the proponents of the affective primacy hypothesis, the presentation times of the stimuli may not have been short enough to prevent (at least partial) awareness of the object in part of the participants or trials (Lähteenmäki et al., 2015; Pessoa et al., 2005). Studies using more sensitive psychometric procedures for threshold detection, and thus more sensitive measures of object recognition, found indeed that physiological responses (e.g., Codispoti et al., 2009; Peira et al., 2012) and reliable judgments of emotion (Peira et al., 2012) as well as reliable valence judgments (pleasant or unpleasant; Lähteenmäki et al., 2015) were only obtained at stimulus presentation times that allowed at least partial object recognition. In fact, according to these findings, not just object recognition (semantic categorization of the eliciting objects) is required for affect but even conscious object recognition (for further evidence, see Hedger et al., 2015).

2.2.2 Chronometric Studies

Because causes must precede their effects (e.g., Hume, 1749/2009), the recognition of an object must precede the elicited affective reaction if the semantic primacy hypothesis holds. Measuring and comparing the onset times of mental events is the purpose of chronometric studies. Posner (1986) defines mental chronometry as "the study of the time course of information processing in the human nervous system" (p. 7). Chronometric methods therefore yield information about the temporal order and temporal distance of mental events, i.e., in our case, object recognition and affect. The semantic primacy hypothesis would be supported if the latency of object recognition is smaller than the latency of affect. If the opposite is found, i.e., if the latency of affect is smaller than the latency of object recognition, the semantic primacy hypothesis is refuted and – in the absence of a plausible third alternative (that is in fact not in sight) – the affective primacy hypothesis is supported. Note that our tests of the semantic primacy hypothesis only require information about the *temporal order* of object recognition and affect onset and the interindividual consistency between the two latencies, not about the size of the delay between the two. However, the temporal judgment experiments reported below also provide information about the size of the cognition-affect delay.

Chronometric Methods for Timing Subjective Events

The first use of chronometric methods to test the semantic versus affective primacy hypothesis dates back to Taizo Nakashima (1909a, 1909b), a student of Edward B. Titchener. In several of his studies, Nakashima measured the latencies of object recognition and affect using speeded reaction time tasks; in additional studies, he also measured the latency of affect with a self-devised temporal judgment method.

Before proceeding, note that a necessary precondition for a meaningful test of the semantic primacy hypothesis is that both compared mental events (i.e., object recognition and affect) are reliably elicited by the stimuli used. Well-suited stimuli that fulfill these requirements in a laboratory setting are pictures showing normatively pleasant (e.g., a laughing child) or unpleasant objects (e.g., maggots); or in the acoustic domain, pleasant (e.g., birds chirping) or unpleasant sounds (e.g., a chainsaw). These stimuli will evoke the normatively typical feelings in the great majority of the participants, although not necessarily in all. Therefore, it is always advisable to check the affect-inducing effects of the stimuli for each participant in an experiment, as we in fact did in our experiments.

Temporal Judgments. In temporal judgment tasks, participants are asked to relate the occurrence of the mental event of interest (e.g., the onset of affect) to that of a temporally close comparison event. This is achieved by presenting the target stimulus – a stimulus that reliably elicits the mental state of interest – and a probe stimulus (which can be in the same or a different stimulus modality) at different short time intervals (stimulus onset asynchronies, SOA) before and after the target stimulus.

In the case of *temporal order* (TOJ) and *simultaneity* (SJ) *judgments*, the probe event consists of a single, easily identifiable stimulus that precedes and follows the target stimulus at different SOAs. In the TOJ task, the participants are asked to report whether the target or the probe event occurred first, or whether the target event occurred before or after the probe event, respectively (Sternberg & Knoll, 1973). In the SJ task, the participants report whether the target and the probe event occurred simultaneously or not (e.g., García-Pérez & Alcalá-Quintana, 2012; Sternberg & Knoll, 1973; van Eijk et al., 2008). In the standard TOJ/SJ procedure, the target stimulus is presented multiple times at each SOA and equally often at each SOA (i.e., the psychophysical method of constant stimuli is used). The data provided by each task are a set of response proportions, one for each SOA: The proportion of “target first” (TOJ), or the proportion of “simultaneous” (SJ) judgments. These data are then used to estimate the *point of subjective simultaneity* (PSS) that reflects the time delay between the onset of the target and the probe stimulus at which the mental events elicited by the stimuli are experienced as simultaneous. While there are several methods for estimating the PSS, in recent years, the standard procedure has become fitting a psychophysical function (a psychometric curve) to the individual proportions data (e.g., García-Pérez & Alcalá-Quintana, 2012; Linares & Holcombe, 2014; van Eijk et al., 2008). Typically, a probit function is fitted to the TOJ and a Gaussian function to the SJ data.

If one sets the onset time of the target stimulus at 0 ms, the PSS directly reflects the delay between the onset of the target mental event and the perception of the probe. Therefore, the PSS provides a (relative) latency estimate for the mental event of interest.

TOJ and SJ tasks have been used in numerous studies to estimate the onset times of subjective events with the aim of investigating a variety of psychological questions for which onset times are relevant. These include the mechanism of multisensory integration (Diederich & Colonius, 2015), the influence of causal beliefs on the perceived timing of action effects (Haering & Kiesel, 2012), and many other – mostly perceptual – phenomena (Jaśkowski, 2014; Sternberg & Knoll, 1973).

The second main kind of temporal judgment task is the *rotating spot method* (e.g., Geiger, 1903; Libet et al., 1983; Pockett & Miller, 2007; Weiß et al., 2013). In this task, the probe event is a fast stimulus stream rather than a single stimulus. In the classical implementation of the method, the stimulus stream is generated by a fast-moving pendulum (Wundt, 1897) or a rotating clock hand (e.g., Geiger, 1903). In modern implementations, the stimulus stream has often been realized by a rotating spot (Libet et al., 1983; Pockett & Miller, 2007), but alternative implementations such as rapidly changing colors (Kosovicheva & Bex, 2020) or a stream of letters (Soon et al., 2008) have also been used. Regardless of the specific implementation of the rotating spot method, it typically looks as follows (in experiments where the target event is induced by an external stimulus): In each trial, the clock (or spot) is started and begins to move. After a shorter or longer delay, the target stimulus is presented. The participants are asked to recall the position of the clock (or spot) at the time the target event occurred. At the end of the trial, they report the remembered position of the clock, for example by indicating it on the clock hand. The reported "clock times" can subsequently be analyzed by methods analogous to those used to analyze reaction times (see below). Pockett and Miller (2007), in a parametric study of the rotating spot method, recommend the use of robust statistics (specifically, the 20% trimmed mean of the latencies) to estimate the latency of the target event separately for each participant, and we followed this advice in our rotating spot studies.

The rotating spot method has been used extensively in introspective psychology (mostly for simple sensations; e.g., Pockett & Miller, 2007; Spence et al., 2001) and has in more recent times become more widely known after Libet et al. (1983) used it to investigate the temporal relationship between the intention to perform a movement and the occurrence of the associated motor readiness potential in the electroencephalogram (EEG). The rotating spot method has, however, also been used in recent times to study temporal relationships in several other areas of psychology (e.g., Carlson et al., 2006; Kosovicheva & Bex, 2020; Seifried et al., 2010; Weiß et al., 2013).

Speeded Reaction Time Task. The speeded reaction time task requires participants to respond as quickly as possible, e.g., by pressing a button or by making a vocal response, when the target mental event occurs (in our case, when the object shown in an affective picture is recognized, and when it elicits a pleasant or unpleasant feeling). Because the distributions of reaction times are typically right-skewed and contain long-latency outliers, a robust estimate of the latency of the mental events is recommended (e.g., Wilcox, 2016), such as the median of the reaction times for that event across multiple stimulus presentations (e.g., Nummenmaa et al., 2010).

Findings of Chronometric Studies

The number of chronometric studies that tested the semantic primacy hypothesis (by comparing the onsets of object recognition and affect) is much smaller than the number of studies on subliminal perception of affective stimuli. The first studies of this kind were, as mentioned, conducted by Nakashima (1909a, 1909b), who used speeded reaction time tasks in several studies to measure both object recognition and affect latency; as well as, in additional studies, a self-devised temporal judgment task to measure affect onset. This method can be reconstructed as a variant of a TOJ task (Nakashima called it the "direct reaction method"). Nakashima used mostly simple stimuli to induce affect, such as color fields, tones, and touch stimuli, but one experiment was conducted with affective pictures, as was done in our experiments. Nakashima found that the "cognitive judgments" took consistently less time than the judgments of affect onset. However, apart from the small number of participants, Nakashima's studies have a number of shortcomings when judged with today's methodological standards. For example, as already mentioned, he only measured the object recognition latencies with the reaction time method, whereas the "direct reaction" (TOJ-type) method was only used for estimating the latency of affect onset. A possible reason may have been that the method is not sensitive enough to detect small latencies, as found in a bachelor's thesis supervised by the author (Mitschka, 2021). Several other shortcomings of the "direct reaction method" are described in Publication A (Reisenzein & Franikowski, 2022).

After Nakashima's pioneering research, chronometric studies on the semantic primacy hypothesis seem to have been largely abandoned until very recently when Nummenmaa et al. (2010) used a speeded reaction time task to compare the latencies of a semantic classification of stimuli (e.g., does a scene contain an animal or a human) to what they called an "affective categorization," that is, judging the stimuli as pleasant or not, or unpleasant or not. Supporting semantic primacy, the authors consistently found, across a set of six experiments, that the median reaction times of the semantic judgments were shorter than those of the affective judgments. It should be noted, however, that Nummenmaa et al. (2010) compared the latency of object recognition to the latency of the detection of stimulus valence, rather than that of the onset of affect. That is, different from Nakashima (1909a, 1909b), they asked their participants to indicate when they judged an object to be pleasant or unpleasant, rather than asking them when they experienced the pleasant or unpleasant feeling (on the importance of this difference, see already Wells, 1925, 1929).

In addition to these chronometric studies of the latency of subjective states, the semantic primacy hypothesis is indirectly supported by studies on the latency of the affective modulation of the late positive potential (LPP), which can be considered the most reliable affect-sensitive component of the event-related potential (ERP; e.g., Codispoti et al., 2006; Cuthbert et al., 2000; Hajcak & Foti, 2020). The LPP is a sustained positive deflection in the ERP waveform that begins 300–400 ms after the onset of a stimulus. This component has been found to be larger in response to pleasant and unpleasant stimuli than to neutral stimuli. This so-called affective modulation of the LPP is thought to be a neurophysiological marker of the attentive processing of motivationally relevant stimuli (e.g., Cuthbert et al., 2000; Hajcak & Foti, 2020). Although it cannot therefore be regarded as a direct marker of the onset of pleasant or unpleasant feelings elicited by affective stimuli, it is intuitively plausible that the onset of feelings and the devotion of increased attention to these stimuli should be temporally close. This assumption is, furthermore, consistent with several theoretical models of emotion (e.g., Bradley, 2009; Reisenzein, 2009). If one accepts this, one can regard the latency of the affective modulation of the LPP as an approximate neurophysiological indicator of the latency of affect, which accordingly is 300–400 ms after stimulus onset. In contrast, both speeded reaction time tasks and ERP studies suggest that *object recognition* can occur within 200 ms (e.g., Fabre-Thorpe, 2011; Förster et al., 2020; Johnson & Olshausen, 2003). Thus, the LPP data also support the semantic primacy hypothesis.

Semantic (Stimulus) Valence Detection Versus Affect Onset

As mentioned, some of the few previous chronometric studies relevant to the semantic versus affective primacy hypothesis did not ask the participants to indicate the onset of a pleasant or unpleasant feeling, but instead asked them to judge the valence of the presented stimulus (Nummenmaa et al., 2010; Wells, 1925, 1929). Judgments about stimulus valence can be viewed as a special kind of evaluative judgment in which a particular evaluative property – hedonic goodness or badness – is attributed to objects. According to a plausible analysis, the valence of an object (for a person) is a dispositional property of the object (e.g., Meinong, 1984; Russell, 2003) that can be defined as the capacity or tendency of the object to evoke pleasant or unpleasant feelings (in that person) under appropriate circumstances (roughly, when the object is presented to the person and the person is in "normal" condition). The detection of this evaluative property of an object, i.e., detecting whether an object is pleasant (capable of evoking a pleasant feeling) or unpleasant (capable of evoking an unpleasant feeling) will here be called "valence detection."

Rather than measuring affect latency (i.e., the latency of the onset of affect), Wells (1925, 1929) and Nummenmaa et al. (2010) thus actually measured the latency of valence detection when they asked their participants to indicate when they rated an object as pleasant or unpleasant. However, valence detection latency can be considered a proxy for the latency of pleasant or unpleasant affect if one assumes that the evaluative judgment is based on the feeling elicited by the object. This is exactly what Wells (1925, 1929) assumed when he used the valence judgments instead of affect onset judgments in his chronometric studies. However, as Wells (1925, 1929) self-critically noted, it is by no means clear that valence judgments are always based on experienced feelings. The reason is obvious: Once the valence of an object is known, this knowledge can be stored in memory just like any other object feature and this knowledge can later be retrieved to make a valence judgment. However, if a valence judgment is based on memory rather than on a feeling, it can no longer be regarded as a proxy for affect and its latency is not diagnostic of the latency of affect. Rather, the latency of a memory-based valence judgment is likely to be similar to the latency of the semantic categorization of the object, i.e., the object recognition judgment.

Two important factors that determine whether valence detection is feeling-based or memory-based are plausibly, (1) the person's experience with the stimulus and (2) the response instruction. For novel objects, hedonic evaluation may depend primarily on the feeling that the object elicits. Consulting one's feelings has been said to be the epistemically primary way to learn the valence of the object (Meinong, 1984) and it may be the usual way to determine the pleasantness or unpleasantness of novel objects (Clore & Byrne, 1974). But even if an object is familiar (e.g., because it has been presented several times during an experiment), valence judgments could still be feeling-based if the stimuli are presented long enough (e.g., for 2,000 ms or more, Franikowski et al., 2021; or until a response is given in a speeded reaction time task, Franikowski & Reisenzein, 2022). If a valence judgment is feeling-based, its latency should be close to that of the latency of affect, because the affect must be registered before it can be used to judge the object's pleasantness or unpleasantness.

For well-known objects, on the other hand, the valence judgment is more likely memory-based. In the simplest case, this means that the perceiver retrieves the valence information stored in the object's memory schema (e.g., Fazio, 2007; Forgas, 1995; Itkes et al., 2017). Alternatively, the valence of the object could be inferred from the category to which it belongs (e.g., "It is always a pleasure to see a pug; therefore, Kevin's pug is pleasant, too.") or from the stored valences of the object's features or components ("The pug's curly tail is cute."). Moreover, memory-based evaluations, if available, are probably more likely if the stimuli are presented under a "fast response" instruction as used in speeded reaction time tasks, because the response can be speeded up by relying on stored valence information rather than on the feeling evoked by the stimulus. In both cases (familiar objects and / or "fast" response instruction), valence judgments could therefore be based on memory-stored valence information rather than on the feeling elicited by the object; and in these cases, the latency of valence detection should be close to that of object recognition.

Some support for these hypotheses stems from recent research by Itkes et al. (2017). These authors found that the intensity of self-reported feelings, facial electromyogram (EMG) responses, and heart rate changes elicited by affective pictures decreased after many repetitions of the stimuli, whereas no such habituation was found for valence judgments. These findings are arguably consistent with the finding that there is rapid habituation of peripheral physiological responses and the overall magnitude of LPP to repetitions of affective pictures, whereas the effect of affective modulation remains intact, i.e., the LPP still differs between affective and neutral pictures after many repetitions (Codispoti et al., 2006; Ferrari et al., 2020; Micucci et al., 2020).

Assuming that valence judgments can be based either on the feeling evoked by an object (particularly for novel objects) or on memory-stored valence information (particularly for well-known objects), predictions regarding the latencies of valence judgments for different stimulus sets can be derived. First, if a stimulus set consists only or mainly of novel objects, valence judgments are based mostly on affect, and their latency should therefore be close to the latency of affect. In addition, the valence detection latencies should in this case be strongly correlated with the affect onset latencies. Second, if a stimulus set contains only or mostly well-known objects, the latency of valence judgments should be close to the latency of object recognition. In addition, the latency of valence detection should be positively correlated with the latency of object recognition. Third, if a stimulus set contains both well-known and novel objects (or better-known and lesser-known objects), the aggregate latency of valence detection should be shorter than that of affect, but longer than that of object recognition, and the latencies of valence detection should be correlated with both the latencies of object recognition and affect. This was considered the most plausible case for the picture presentations in our Experiment C2, where this hypothesis was tested.

Motivation and Rationale for Cross-Method Replications

In this dissertation project, a variety of different latency measurement methods was used: temporal order and simultaneity judgments, the rotating spot method, and the speeded reaction time task. The rationale behind using these different chronometric methods was this: Each chronometric measurement method has its strengths and weaknesses, but the biases of different measurement methods are unlikely to be exactly the same. Therefore, if the results obtained with one measurement method can be replicated by another method, this strengthens the conclusions that can be drawn from each method (see Linares & Holcombe, 2014).

What are the advantages and disadvantages of the different chronometric methods? The focus here will be on the contrast between temporal judgments and speeded reaction time measurements.

An advantage of speeded reaction times is that they are easy to collect and provide absolute response latencies, whereas temporal judgments are more effortful to collect and only yield relative latencies. A main disadvantage of speeded reaction time measures is that the obtained latencies are a biased estimate of the latencies of the focal mental events, because response latencies also include (motor) decision and execution time for the response (for discussions, see Fabre-Thorpe, 2011; Jaśkowski, 2014; Miller & Ulrich, 2013; Neumann & Niepel, 2004). Therefore, latencies measured with speeded reaction times can only be considered indirect measures of the perceptual latencies. In addition, the *differences* between the latencies of different mental events tend to be two to four times larger when measured with speeded reaction time tasks than when measured with temporal judgment tasks (Cardoso-Leite & Gorea, 2010; for an explanation of this effect, see Miller & Schwarz, 2006). Although this does not prevent testing hypotheses on temporal order and interindividual consistency of the latencies, it suggests that speeded reaction time tasks do not yield precise temporal difference estimates. While the focus of the dissertation research program was on the temporal order of object recognition and affect and its interindividual consistency (indicated by a positive correlation), for estimating the size of the time difference we therefore relied on the findings obtained with temporal judgments.

In contrast to speeded reaction time tasks, temporal judgment tasks can be assumed to be unbiased by motor decision and execution times, and thus to provide a purer estimate of perceptual latencies (Schneider & Bavelier, 2003). However, temporal judgment tasks also do have some disadvantages: One disadvantage, as already mentioned above, is that temporal judgments only yield relative response latencies (the time differences between a target mental event and the perception of the probe event; Schneider & Bavelier, 2003). This disadvantage is, however, not relevant for testing the semantic primacy hypothesis (and, in fact, for testing most other temporal hypotheses of emotion psychology), because for this purpose, information about the temporal order of perceptual latencies is sufficient. In addition, if the absolute latency of just one of a series of relative latencies obtained in a temporal judgment task is known, then the absolute latencies corresponding to the remaining latencies can be calculated from the latency differences to the reference latency (Reisenzein & Franikowski, 2022).

A second disadvantage of temporal judgment tasks is the possibility of dual-task interference, because subjects must simultaneously focus on the appearance of the probe event and the target event. However, as long as the dual-task interference effect is additive in nature (i.e., the latencies of object recognition and affect are equally prolonged), estimates of temporal difference are unbiased. In contrast, if the dual-task interference effect varies for different mental events (i.e., if the latencies of object recognition and affect onset are affected differently), the estimates of the temporal difference between the events would be systematically biased and even the test of the temporal order hypothesis can be jeopardized. Because the dual-task problem is particularly salient in the rotating spot method (e.g., Danquah et al., 2008; Verbaarschot et al., 2016), the consequences of this interaction and its implications are thoroughly discussed in Publication B (Franikowski et al., 2021).

Experimental Extensions of the Chronometric Tests of Semantic Primacy

Measuring, comparing, and correlating the estimated latencies of object recognition and affect provides information about the temporal order of the latencies of object recognition and affect and their interindividual consistency. If the order is as predicted and the correlation between the latencies is positive, this supports the semantic primacy hypothesis. Yet, this support is only limited. Recall that the semantic primacy hypothesis (as understood in this research program) is a causal hypothesis, i.e., it posits that object recognition is a necessary partial cause of affect. Temporal priority of object recognition, and a positive correlation between object recognition and affect latency are necessary for object recognition being a cause of affect but not sufficient. The strongest and most convincing test of a causal hypothesis is an experimental test in which the assumed cause is experimentally manipulated and its effect on the hypothesized dependent variable is measured.

Following this logic, in Publications B (Experiment B2) and C (Experiment C2), the onset of object recognition was experimentally delayed by a moderate blurring of the affective pictures. The semantic primacy hypothesis entails that any manipulation that delays object recognition will also delay the onset of affect. In addition, it predicts that the effect of blurring on the latency of affect will be statistically mediated by delayed object recognition. If these predictions are supported, that would provide strong evidence for the semantic primacy hypothesis.

The latency of affect is certainly also determined by factors in addition to the occurrence time of object recognition. Therefore, there should also be experimental manipulations that delay only affect but not object recognition. Such manipulations could be, for example, manipulations that reduce the intensity of the experienced affect (e.g., by altering the evaluation of the stimulus) without influencing object recognition latency. The existence of such effects does, however, not refute the semantic primacy hypothesis, because according to this hypothesis, object recognition is only a partial cause of affect. In one of the experiments reported here (Experiment C2) we did include such a manipulation that (in similar form) was found to strongly reduce the intensity of pleasant and unpleasant feelings evoked by pictures in a previous study (Junge & Reisenzein, 2013): false-coloring affective pictures.

2.3 Hypotheses

A series of eight experiments (already published in three journal articles) was designed to test the semantic primacy hypothesis. The first four experiments, reported in

Publication A[2] (Reisenzein & Franikowski, 2022), used temporal order judgments (Experiments A1a, A1b, and A3) and simultaneity judgments (Experiment A2). In Experiments B1 and B2, reported in Publication B (Franikowski et al., 2021), the rotating spot method was used to measure the perceptual latencies and in Experiments C1 and C2 of Publication C (Franikowski & Reisenzein, 2022), the speeded reaction time method was used.

In all experiments, two central predictions of the semantic primacy hypothesis[3] were tested. The first one was:

H1: The latency of object recognition is shorter than the latency of affect.

The derivation of this hypothesis from the semantic primacy hypothesis is simple. The only premise additional to the semantic primacy hypothesis is that causes must precede their effects. This premise is generally regarded as valid: It belongs to our very understanding of causality that causes must occur before their effects (Mackie, 1974); causality goes with the flow of time. If object recognition is a partial cause of affect (semantic primacy hypothesis), it must therefore occur before the onset of affect. What Hypothesis 1 does not specify is how much later the onset of affect will occur. However, the results of Nummenmaa et al. (2010) suggest that the delay may be in the order of 40–80 ms, and should be detectable even with small participant numbers in a within-subjects design (see Experiment A1 for more detail). The findings of Nakashima (1909a, 1909b) for unambiguous affect onset judgments suggest that the time lag is even greater.

The second central prediction addresses the interindividual consistency of the aforementioned temporal order of object recognition and affect:

H2: There is a positive correlation between the latencies of object recognition and affect.

In Publications B and C, we conducted an experimental test of the semantic primacy hypothesis in the latency measurement paradigm (Experiments B2 and C2): We manipulated the recognizability of the objects shown in the affective pictures by moderately blurring one half of the picture set. Moderate blurring was used because it was important that the manipulation did not completely prevent object recognition, but only delayed it in comparison to the unaltered (sharp) presentation of the pictures (De Cesarei & Codispoti, 2010; Schupp et al., 2008). The hypotheses for this manipulation were:

H3: Blurring of (pictorial) affective stimuli delays not only object recognition but also the onset of affect.

[2]Although published after the Publications B and C, the Experiments A1a, A1b, and A2 were conducted first. Experiment A3 was added during the review process to answer objections by the reviewers but used the same method (temporal order judgments) as Experiments A1a and A1b.

[3]In all experiments (Franikowski et al., 2021; Franikowski & Reisenzein, 2022; Reisenzein & Franikowski, 2022), several additional hypotheses were tested. These concerned the temporal relation between, on the one hand, the perceptual latencies of object recognition and affect, and on the other hand, the latencies of various control judgments that were assumed to be simpler than these: the detection of stimulus appearance, color, and shape. To not digress from the focal hypotheses, the results obtained for these additional hypotheses are ignored here but they are reported at the beginning of the Summary section 3.1 and in the published articles.

H4: The delaying effect of blurring on the onset of affect is statistically mediated by the delay in object recognition.

Because only partial mediation was found in our first experimental test of the semantic primacy hypothesis in Experiment B2 (Publication B), and because the same study yielded preliminary but inconclusive evidence that affect intensity may also have been involved in mediating the effect of blurring on delayed affect latency, we also wanted to clarify the effect of affect intensity on affect latency in Experiment C2 (Publication C). To achieve this goal, we included a second experimental manipulation in this study that was designed to reduce affect intensity without influencing the latency of object recognition. We decided to use false-coloring the pictures for this purpose because a very similar manipulation had been found to strongly reduce affect intensity in a previous study (Junge & Reisenzein, 2013, Experiment 2). At the same time, this manipulation left the coarser and finer contours of the objects largely untouched, and it is on these that object recognition mainly depends (see Experiment C2). Therefore, the following predictions were made regarding the effects of false-coloring:

H5: False-coloring the affective stimuli delays the onset of affect but not object recognition.

H6: The delaying effect of false-coloring the affective stimuli on the onset of affect is mediated by a decrease in affect intensity.

While the six experiments of Publications A and B measured only the latencies of object recognition and affect, Experiments C1 and C2 (in Publication C) also estimated the time needed for detecting the valence of the stimuli. This was done to compare the latency of valence detection with the latencies of object recognition and affect. As explained above, valence detection was hypothesized to occur later than object recognition but before affect onset (for which it has been regarded as a proxy by, e.g., Nummenmaa et al., 2010; Wells, 1925, 1929), especially for a set of both well-known and novel objects. Furthermore, these temporal order relations are also assumed to be interindividually consistent. In addition, the latency of valence detection was also hypothesized to be affected by the experimental manipulations of blurring and false-coloring in ways parallel to the latency of affect. In Experiments C1 and C2, accordingly, the following additional hypotheses, concerning the latency of valence detection, were tested:

H7: The latency of valence detection is (a) longer than the latency of object recognition and (b) shorter than the latency of affect.

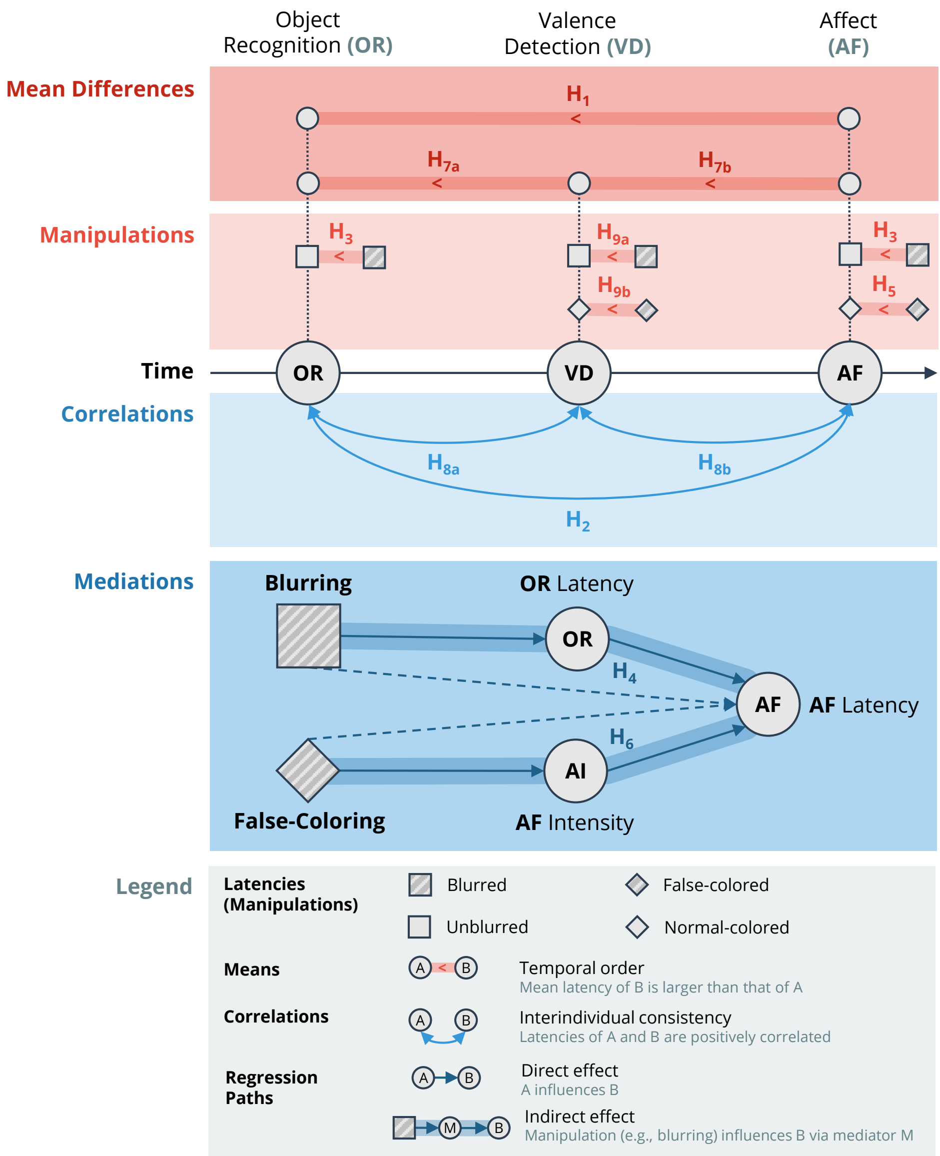
H8: The latency of valence detection is positively correlated with (a) the latency of object recognition and (b) the latency of affect.

H9: The latency of valence detection is delayed by (a) blurring and (b) false-coloring the affective stimuli.

The described hypotheses are summarized graphically in Figure 2.1.

Figure 2.1

Graphical Abstract of Hypotheses



Note: Bold straight lines without arrowheads reflect mean difference (temporal order) hypotheses; curved arrows with two arrowheads indicate correlations; single-headed arrows indicate direct regression effects; bold lines above regression paths represent the indirect (mediated) regression effect. The numbering of the hypotheses used in this figure corresponds to the numbering in the text. Note that the hypotheses are partly numbered differently in the publications.

Chapter 3

Summary of the Experiments

This section summarizes the experiments that tested the described hypotheses. In addition, it includes a meta-analytic integration of the results to profit from the increased power of the meta-analytic integration and to obtain an integrated estimate of the effect sizes for the central Hypotheses 1 (temporal difference) and 2 (correlation).

3.1 Common Features of the Experiments

All experiments tested the semantic primacy hypothesis with pictorial affective stimuli. The affective stimulus sets comprised pictures that were rated in previous studies regarding the intensity of pleasant or unpleasant feelings they elicit (Junge & Reisenzein, 2013; Kurdi et al., 2017). In addition to testing the effects of the central factors task (object recognition vs. affect vs. valence detection; all experiments), blurring (sharp vs. blurred pictures; Experiments B2 and C2), and false-coloring (normal-colored vs. false-colored pictures; Experiment C2), we also checked for possible effects of picture valence and task order by including these as control variables into the analyses in all paradigms (for an overview of all control factors, see Figure 3.1).

In all experiments, the chronometric tasks were preceded by a rating task, i.e., participants were presented with the affective pictures and asked to rate the intensity of pleasant or unpleasant feelings (depending on picture valence according to the ratings from previous studies, Junge & Reisenzein, 2013; Kurdi et al., 2017) evoked by the pictures. In the studies reported in Publication A, we repeated this rating at the end of the experiment to verify that the stimuli had maintained their affect-inducing capacity despite the frequent presentations. Because this was found to be the case, the affect ratings in the subsequent experiments (see Publications B and C) were made only at the beginning.

Judged by the affect ratings, the pictures successfully elicited at least moderately pleasant or unpleasant affect in all but a few participants in all studies. If the pictures failed to induce at least moderately pleasant or unpleasant affect in a participant, they were excluded from the main analyses because it is not meaningful to measure the onset of a feeling if the feeling does not occur. However, the inclusion of these (few) participants in a total analysis did not change the results or their interpretations.

The first chronometric task block (or blocks) always consisted of one or more training tasks intended to be simpler than the focal tasks (object recognition, affect, and valence detection). These simpler tasks served to familiarize the participants with the chronometric procedure and to validate the measurement method. Specifically, the participants were asked to detect the appearance of a simple stimulus (e.g., a red square; Experiments B1 and B2), and / or to detect the color of a uniformly colored field (e.g., a blue field; all experiments except for Experiment B2) or the shape of a simple geometric figure (e.g., a circle; Experiment B2). We found, as predicted, that the latency of appearance detection was smaller than that of color (and shape) recognition, and that the latency of color (and shape) recognition was smaller than the latencies of object recognition and affect. This was found for almost all participants.

Following the training tasks, the participants were asked to judge, in separate blocks of trials, when they recognized the object shown in the picture (object recognition) and when they experienced the feeling elicited by the picture (affect).

To control for possible order effects, i.e., to rule out the possibility that a positive latency difference between affect and object recognition was due to fatigue or that a negative difference was due to practice, we varied the order of the focal tasks (task order: object recognition vs. affect first) in most experiments (except for Experiments A1b and C1). Depending on the specific experiment and its implementation, additional control factors were included (see Figure 3.1).

The basic procedure for testing the hypotheses about mean differences was identical in all experiments: In the first step, an analysis of variance (ANOVA) model was estimated that included the factors of focal interest plus the control factors. In the second step, control factors without a main effect or an interaction effect with any of the focal factors were omitted, and the results obtained for this reduced ANOVA model were interpreted (see Figure 3.1).

The sample sizes for all experiments were determined based on power analyses for the central Hypotheses 1 (temporal difference) and 2 (correlation). While the power analyses of the first set of experiments (Publication A) was based on the findings of Nakashima (1909a, 1909b) and Nummenmaa et al. (2010), the power analyses for the other experiments were based on the results of the previous experiments in Publication A (for Publication B) or Publications A and B (for Publication C).

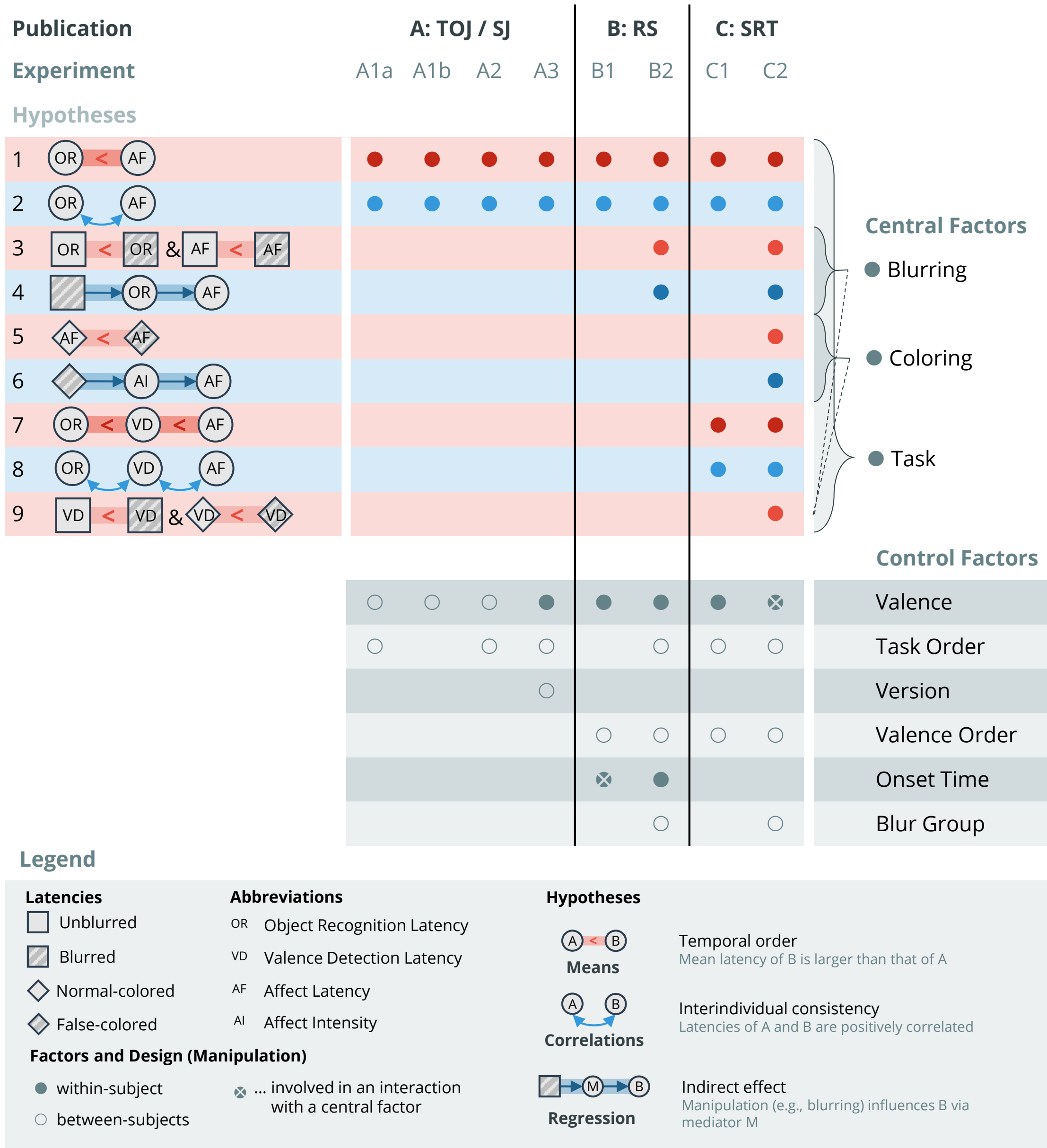
3.2 Publication A: Temporal Order (TOJ) and Simultaneity Judgments (SJ)

In the first set of studies, reported in Publication A (Experiments A1–A3), we tested the semantic primacy hypothesis using temporal judgments, more precisely temporal order judgments (TOJ; Experiment A1a, A1b, and A3) and simultaneity judgments (SJ; Experiment A2).

In all four experiments, pictorial affective stimuli were presented that were either preceded or followed by a visual probe consisting of two blue bars that appeared above and below the picture with varying stimulus onset asynchronies (SOAs; see Figure 1 in Publication A, Reisenzein & Franikowski, 2022). Depending on the method (TOJ or SJ), the participants had to respond to a TOJ ("Did object recognition [affect] occur before or after the probe?") or a SJ question ("Did object recognition [affect onset] occur simultaneously or not simultaneously with the probe?"). The participants' responses were accumulated into relative

Figure 3.1

Overview of Central and Control Factors Across the Eight Experiments



Note: TOJ = temporal order judgments; SJ = simultaneity judgments; RS = rotating spot task; SRT = speeded reaction time task. A filled circle indicates that the factor was manipulated within-subject, whereas an empty circle indicates that the factor was manipulated between-subjects. Crossed-out circles indicate that a control factor was involved in a significant interaction with one or more of the central factors and was therefore also included in the main analyses of the respective hypotheses. For details, readers are referred to the published articles.

frequencies of “before” (TOJ) or “simultaneous” (SJ) judgments for the different SOAs, and subject-level psychometric curves were fitted to these frequencies to determine each individual’s point of subjective simultaneity (PSS) for both tasks (object recognition and affect). The PSS provides an estimate of the relative latencies of object recognition and affect (relative to the onset of the perception of the probe). The fit of the psychometric curves was assessed using R^2. Participants were excluded based on low affect ratings (general criterion) and bad fits of the psychometric curves (implausible response patterns and / or low R^2; paradigm-specific criterion). However, curve fit was very good to excellent for almost all participants in the four experiments.

According to Hypothesis 1, the PSS of object recognition should be shorter than the PSS of affect. Furthermore, according to Hypothesis 2, both latencies should be positively correlated. Confirmation of these hypotheses supports the semantic primacy hypothesis.

3.2.1 Experiment A1a: TOJ

Method

We recruited $N = 20$ undergraduate students as participants. The participants were presented with either 12 pleasant or 12 unpleasant (disgusting) stimuli. Each picture was presented 12 times per task block (object recognition vs. affect) with each presentation at one of 12 different SOAs (between -746 and +746 ms). This resulted in 144 presentations per task. In addition, picture valence (pleasant vs. unpleasant) and task order (object recognition vs. affect first) were included as between-subjects control factors, resulting together in a 2 × 2 between-subjects design. Together with the within-subject factor task (object recognition vs. affect), the experiment thus had a 2 × 2 (between) × 2 (within) design.

Results

Two participants were excluded from further analyses because they had very low ratings of affect intensity. However, including these subjects into the analyses did not change the results or their interpretation. We found that the average difference between the latencies of affect and object recognition was significant and positive with an average delay of 79 ms. This supports Hypothesis 1. Furthermore, the latency difference was positive in 15 of the 18 cases. The correlation between object recognition and affect latencies was positive but not significant. Therefore, Hypothesis 2 received only limited support.

3.2.2 Experiment A1b: TOJ With a Broader SOA Range

Method

Experiment A1b served as a control experiment for Experiment A1a. The psychometric curve of one participant of Experiment A1a suggested that their PSS for affect onset may have exceeded the upper SOA of +746 ms. As a consequence, a much wider SOA window ranging from -1,800 and +1,800 ms was used in Experiment A1b for two reasons: First, to account for the possibility that the onset of affect for some individuals is indeed later than +746 ms. Second, to rule out the possibility that the participants of Experiment A1a felt obliged to report faster onset times because they inferred from the provided SOA range that affect onset was expected to occur before +746 ms.

The latencies of object recognition and affect were measured with TOJs in N = 13 participants from the same subject pool as in Experiment A1a (undergraduate students). Again, either 12 pleasant or 12 unpleasant stimuli (the same as in Experiment A1a) were shown. Also, as in Experiment A1a, each picture was presented with the probe displaced by 12 different SOAs.

Because this control experiment was designed specifically to examine whether the SOA range used in Experiment A1a was broad enough, a smaller sample size was used. In Experiment A1a, task order did not have a significant main effect or interaction effect with the central factor task or the control factor picture valence. Therefore, task order was not varied, but the affect onset judgments were always made second. However, we controlled again for valence effects by manipulating the valence factor between-subjects.

Results

The ANOVA showed that the average latency difference between affect and object recognition was again significant and positive (113 ms). Furthermore, the latency difference was positive for 11 of the 13 participants. The correlation between the latencies of object recognition and affect was positive but again not significant. Therefore, Hypothesis 2 (correlation) received again at best limited support, whereas Hypothesis 1 (latency difference) received full support.

3.2.3 Experiment A2: SJ

Method

The procedure of the third experiment was again parallel to that of Experiment A1a, but with the TOJ task exchanged for the SJ task. In parallel with Experiment A1a, we had a 2 (task order) × 2 (valence) between-subjects design. In an attempt to improve the accuracy of the PSS estimation of object recognition, we reduced the SOA range for object recognition to between -200 and +200 ms (instead of -746 and +746 ms). This SOA range included the locations of the individual PSSs in the first experiment. We collected data from N = 20 undergraduate students. One individual was excluded from further analyses because of low ratings of affect intensity. Three other participants were excluded from the analyses because of implausible psychometric curves; the curves either indicated too high error rates or the PSSs for affect could not be clearly identified within the given SOA range.

Results

We found a significant main effect of task reflecting again a later PSS for affect than for object recognition. The latency difference averaged 154 ms and the difference was positive for all subjects. These findings clearly support Hypothesis 1. The correlation was also positive but not significant, providing limited support for Hypothesis 2.

3.2.4 Experiment A3: Generalization

Method

Experiments A1a, A1b, and A2 had at least two limitations. First, they included only a small and select sample of stimuli; it is therefore possible that the results would not generalize to other affective stimuli. This possibility is particularly salient for the negative stimuli because only disgust-eliciting pictures were used. Second, because pleasant and unpleasant pictures were judged in separate blocks (this was done to facilitate the task), it is conceivable that there was an accumulation of affect within each valence-homogeneous block. As a consequence, it could have become increasingly difficult for the participants to identify the onset of affect elicited by a picture against the background of the rising baseline. If so, this may have artifactually prolonged affect latency.

To examine the validity of these objections, Experiment A3 was devised. This experiment was conducted with $N = 40$ participants (mostly undergraduate students) and was a modified replication of Experiment A1a. To make a stronger case for stimulus generalization, we presented 168 pictures (84 pleasant *and* 84 unpleasant) instead of only 12 pictures (12 pleasant *or* 12 unpleasant) as in the preceding experiments. The pictures were selected from the free OASIS picture pool (Kurdi et al., 2017). They consisted of pictures from widely different object categories that had high normative ratings of pleasure and displeasure, respectively, and were rated as eliciting at least moderate affect in a pretest ($N = 20$ students rated these OASIS pictures in a separate online study). In addition, two SOAs were added to the 12 SOAs from Experiment A1a (-1,200 and 1,200 ms) to improve estimation of the psychometric curve. Each picture was randomly presented at one SOA, while each SOA was represented by 12 picture presentations (6 pleasant and 6 unpleasant, respectively). The assignment of pictures to their respective SOAs was randomly generated for each subject but remained identical for both tasks (object recognition and affect). In contrast to the previous experiments, pleasant and unpleasant pictures were presented in a randomly intermixed sequence to protect against the possibility of affect accumulation.

Because of the difficulties of recruiting participants for laboratory studies during the COVID-19 pandemic in 2021, the experiment was conducted as a laboratory version and a web version (factor experiment version). We also manipulated task order between-subjects in parallel with Experiments A1a and A2, resulting in a 2 (experiment version) $\times$ 2 (task order) between-subjects design.

We finally analyzed the data of $N = 34$ subjects. For four of the 40 participants, the data were incomplete because of Internet connection problems; two participants were excluded because of a poor fit of the psychometric curves. Conducting the analyses without excluding the latter two participants did not significantly alter the results or their interpretation.

Results

The PSS difference for affect and object recognition was significantly positive and averaged 78 ms. Twenty-six of the 34 participants had a positive latency difference. The correlation between the latencies of object recognition and affect was again positive and, in contrast to the preceding experiments, was significantly different from

0. In Experiment A3, then, both Hypotheses 1 (mean difference) and 2 (correlation) were fully confirmed.

3.2.5 Discussion of Experiment Series A (Publication A)

All four experiments (as well as their meta-analytic integration; see Publication A, Reisenzein & Franikowski, 2022) provide evidence for the semantic primacy hypothesis. We found the predicted positive latency difference between affect and object recognition (H1) both at the group level and at the individual level for almost all participants in all four experiments. In addition, we obtained a positive correlation between the two latencies (H2) in all experiments, but this correlation was significant only in Experiment A3. Neither the task order nor the valence of the affective pictures (nor, in Experiment A3, whether the study was conducted in the lab or online) affected the results in terms of interactions; the obtained latency difference therefore cannot be attributed to these factors.

3.3 Publication B: Measuring Latencies With the Rotating Spot Method

In the second series of experiments, reported in Publication B (Experiments B1 and B2), we investigated semantic primacy by measuring the latencies of object recognition and affect with the second main kind of temporal judgment methods, the rotating spot paradigm (e.g., Libet et al., 1983; Pockett & Miller, 2007).

The two experiments used almost the same stimuli as in Experiments A1a, A1b, and A2. In each trial of the rotating spot task, participants first saw a gray picture surrounded by a blue frame symbolizing a clock face. Shortly thereafter, a rotating spot symbolizing the clock hand appeared and began to rotate around the picture, lingering at each of 20 consecutive positions for 107 ms. Hence, one rotation lasted 2,140 ms. After a variable interval after the start of the clock, an affective picture was shown. In different blocks of trials, the participants had to remember and later report the position of the rotating spot when the mental event in question occurred (focal tasks: object recognition and affect; see Figure 1 in Publication B, Franikowski et al., 2021). To estimate the latencies of object recognition and affect, we calculated the 20% trimmed mean, following a recommendation of Pockett and Miller (2007). This recommendation is based on the fact that the "clock times" reported in the rotating spot task typically have a right-skewed distribution with potentially large long-latency outliers. The 20% trimmed mean instead of the median (which is actually the 50% trimmed mean) was used because it provides for a more fine-graded response scale.

The aim of the first experiment (B1) was to replicate the results of the studies in Publication A using the rotating spot method, whereas the second experiment (B2) extended these results by adding an experimental manipulation of the onset time of object recognition (a slight blurring of the picture, that delayed object recognition) to provide an experimental test of the semantic primacy hypothesis in the latency measurement paradigm. In addition to Hypotheses 1 (there is a positive latency difference between affect and object recognition) and 2 (the two latencies are positively correlated), Experiment B2 therefore also tested Hypotheses

3 (delaying object recognition also delays the onset of affect) and 4 (the delaying effect of blurring on affect latency is mediated by the delayed object recognition). As in the experiments of Publication A, participants who had very low affect ratings and / or implausible response patterns (see below for examples) were excluded from further analyses.

3.3.1 Experiment B1

Method

For Experiment B1, we recruited $N = 19$ undergraduate students. Prior to data analysis, one participant was removed due to implausible latencies (the latencies of object recognition and affect were shorter than those for detecting the onset of a simple stimulus), reducing the final sample size to $N = 18$. Participants were presented with almost the same pictures as in Experiments A1a, A1b, and A2. In contrast to these studies, all participants were presented with all 12 pleasant and 12 unpleasant pictures (instead of only seeing either pleasant or unpleasant pictures). However, in contrast to Experiment A3, the pleasant and unpleasant pictures were not presented intermixed but in two separate blocks within the focal task blocks. To control for potential order effects of valence blocks, we manipulated the valence order (pleasant vs. unpleasant pictures first) between-subjects. To prevent participants from developing expectations about the onset of the pictures, each picture was presented four times per task block, once at each of four different onset positions (i.e., 107, 214, 321, or 428 ms after rotation onset). Taking the central design factors and the control factors together, we obtain a 2 (task) $\times$ 2 (valence) $\times$ 4 (picture onset times) within-subject $\times$ 2 (valence order) between-subjects design.

Results

Because picture valence and valence order had no significant main or interaction effect on the latencies, these factors were excluded from further analyses. However, because the main effect of task and its interaction effect with (and the main effect of) the control factor picture onset time were significant, we added it to the final simplified 2 (task) $\times$ 4 (picture onset times) ANOVA. We found a significant main effect of task, which was also confirmed by simple-effects ANOVAs for each of the four levels of the control factor picture onset time. The average latency difference between affect and object recognition was positive for all 18 participants, averaging 205 ms. Therefore, Hypothesis 1 was fully supported. Confirming Hypothesis 2, we also found a significant positive correlation between the latencies of object recognition and affect.

3.3.2 Experiment B2: Experimental Manipulation (Blurring)

Method

For the second experiment, we recruited $N = 40$ participants (mostly undergraduate students), two of whom were excluded due to extremely deviant affect latencies, resulting in a final sample size of $N = 38$. The procedure of Experiment B2 was parallel to Experiment B1 except for the following, minor changes: (1) A few unpleasant pictures that were very difficult to recognize in the blurred condition were exchanged for other pictures that were similarly well recognizable in blurred form as the other blurred pictures. (2) The pleasant and unpleasant pictures were split into parallel halves (set 1 and set 2) based on ratings of the pictures' affect intensity in Experiment B1. One half of the participants saw set 1 in the sharp and set 2 in the blurred version, the other half saw set 2 in the sharp and set 1 in the blurred version. That way, we ensured that no participant saw a blurred picture in its sharp version, since that might have facilitated recognition in later trials. To verify that there were no systematic differences between the two picture sets, we included blur group (first vs. second half blurred) as a control factor into the ANOVAs. (3) Task order was again varied as in Experiments A1a, A2, and A3. The inclusion of blur group and task order meant that the 2 (task) × 2 (valence) × 4 (picture onset times) within-subject × 2 (valence order) between-subjects design of Experiment B1 was extended to include another within-subject factor of focal interest, blurring (sharp vs. blurred) and two additional between-subjects control factors with 2 levels each (task order; blur group).

Results

A 2 (blurring) × 2 (valence) within-subject × 2 (blur group) between-subjects ANOVA on affect ratings revealed that only the main effects of blurring and valence were significant. As intended, blurring reduced affect intensity. In addition, pleasant pictures elicited slightly more intense affect than unpleasant pictures.

A preliminary ANOVA of the latencies, with all focal and control factors included, showed significant effects of task, blurring, and picture onset time. No other main or interaction effect of the other control factors was significant. Therefore, the other control factors were dropped and a simplified 2 (task) × 2 (blurring) × 4 (picture onset time) ANOVA was computed. Only the main effects of the three factors were significant in this ANOVA. In addition, the main effect of task was significant in subsequent simple effects analyses for all four picture onset times and the two picture conditions (sharp vs. blurred). The latency difference between affect and object recognition was positive for 31 of the 38 participants and averaged 124 ms. Therefore, Hypothesis 1 was supported in Experiment 2. In addition, the correlation between the latencies of object recognition and affect was positive and significant, fully supporting Hypothesis 2.

Most importantly, the experimental test supported the semantic primacy hypothesis: First, in support of Hypothesis 3, we found that blurring the pictures caused a significant delay of both object recognition (showing that the intended manipulation was successful) and affect. Second, partially supporting Hypothesis 4, a multilevel path analysis showed that the delaying effect of blurring on affect latency was in part mediated by the delay in object recognition latency.

3.3.3 Discussion of Experiment Series B (Publication B)

The results of both Experiments B1 and B2 supported the semantic primacy hypothesis. Replicating the findings of the temporal order and simultaneity judgment studies (see Publication A), measuring the latencies with the rotating spot method found a positive latency difference between affect and object recognition (H1) and a positive correlation between these latencies (H2); in Experiments B1 and B2, this correlation was also larger than in most experiments of Publication A and was significant in both experiments. Most importantly, Experiment B2 provided for an experimental test of the semantic primacy hypothesis: Blurring the affective pictures was a manipulation that was hypothesized, and found, to delay object recognition. As predicted by Hypothesis 3, this manipulation also delayed the onset of affect. In addition, Hypothesis 4, that the delaying effect of blurring the affective pictures on affect is mediated by the delay in object recognition, received partial support: A mediation analysis revealed that the effect was indeed partly but not completely mediated by the delayed object recognition. Experiment B2 also provided suggestive (but inconclusive) evidence that another mediator of the blurring effect on affect latency was the intensity of affect. That is, blurring also reduced the intensity of the feelings evoked by the affective pictures, and weak feelings probably need more time to be detected (similar to weak sensory stimuli). In Experiment C2, reported below, the hypothesis that a reduction of affect intensity causes a delayed onset of experienced feelings, was examined in more detail.

3.4 Publication C: Speeded Reaction Times

The experiments of Publication A tested Hypotheses 1 (object recognition precedes affect) and 2 (positive correlation of object recognition and affect latency); those of Publication B additionally tested Hypotheses 3 (delaying effect of blurring on object recognition and affect) and 4 (mediation of the delaying effect of blurring on affect via delaying object recognition). In the experiments of Publication C, these four hypotheses were tested again, but using speeded reaction time measurements to estimate perceptual latencies. The participants were asked to press a key on the computer keyboard as quickly as possible when the mental event of interest occurred, i.e., when they recognized what was shown in the picture (object recognition) or when they experienced the feeling elicited by the picture (affect). In addition, we added another dependent variable, the latency of valence detection. In the valence detection task, the participants were asked to press one of two buttons as fast as possible when they recognized that the picture was pleasant or unpleasant, respectively. Four new hypotheses were a priori derived for this latency (Hypotheses 5–9).

The purposes of Experiment C1 were: First, to replicate the results from the temporal judgment studies (Publications A and B) for the basic Hypotheses 1 and 2 with the speeded reaction time task. Second, Experiment C1 allowed to compare the latencies of valence detection to the latencies of object recognition and affect, and to compute the correlations of these latencies. This allowed to test Hypotheses 7 (the latency of valence detection is between the latencies of object recognition and affect) and 8 (the valence detection latency correlates positively with the latencies of object recognition and affect).

In Experiment C2, we repeated the experimental test of the semantic primacy hypothesis conducted in Experiment B2 by including the factor picture blurring. This allowed to retest Hypotheses 3 and 4. In addition, Experiment C2 included a second experimental manipulation, false-coloring of the pictures. This allowed to test Hypotheses 5 (false-coloring delays the onset of affect but not object recognition), 6 (the delaying effect on false-coloring on affect is mediated by reduced affect intensity), and 9 (blurring and false-coloring also delay the detection of valence).

3.4.1 Experiment C1

Method

For Experiment C1, we recruited $N = 20$ undergraduate students. Two participants were excluded due to very low ratings of affect intensity, reducing the sample size to $N = 18$. Participants were presented with the same 12 pleasant and 12 unpleasant pictures as in Experiments A1a, A1b, and A2, which as mentioned were also mostly identical to the pictures used in Experiments B1 and B2. For theoretical reasons (see Publication C, Franikowski & Reisenzein, 2022), the valence detection task was always presented as the last task. To control for possible order effects, the order of the preceding task blocks of object recognition and affect (factor task order) and the order of the valence blocks within these two task blocks (factor valence order) were varied between-subjects. Thus, Experiment C1 had a 3 (task: object recognition vs. valence detection vs. affect) $\times$ 2 (valence) within-subject $\times$ 2 (task order) $\times$ 2 (valence order) between-subjects design. All pictures were presented twice per task $\times$ valence block.

Results

The difference between the latencies of affect and object recognition was significantly positive and averaged 511 ms, supporting Hypothesis 1. Furthermore, this latency difference was found to be positive for all 18 participants. In addition, the latency difference between valence detection and object recognition (H7a) and that between affect and valence detection (H7b) were positive in at least 15 of the 18 participants (and also significantly positive on the overall level, respectively), supporting Hypothesis 7. All predicted correlations between latencies were positive, but none was significant, possibly due to the low sample size. Therefore, there was only limited support for Hypotheses 2, 8a, and 8b.

3.4.2 Experiment C2: Experimental Manipulation (Blurring and False-Coloring)

Method

Forty undergraduate psychology students were recruited for Experiment 2. One participant was excluded from the data analyses because of very low ratings of affect intensity. The procedure was completely parallel to that of Experiment C1, with the exception of the identity of the pictures and the introduction of two experimental manipulations (i.e., blurring and false-coloring). Because one aim of Experiment C2 was to replicate the effects of the blurring manipulation of Experiment B2, we used the same pictures as in Experiment B2 and presented either the first or the second of the two parallel halves of the pictures in a blurred condition. For the false-coloring manipulation, the red-green-blue (RGB) color channels of the pictures were converted to blue-red-green (BRG). Therefore, in Experiment C2, two 2-level within-subject factors of focal interest (blurring; coloring) and one 2-level between-subjects control factor (blur group) were added to the design of Experiment C1. All pictures were presented four times per task $\times$ valence block, twice in the normal-colored and the false-colored condition, respectively.

Results

Affect intensity ratings differed depending on blurring, coloring, and valence (in terms of a three-way interaction). A significant main effect of blurring replicated the result of Experiment B2 in Publication B, i.e., blurring the pictures reduced the intensity of the experienced affect. As predicted, false-coloring the pictures also reduced the intensity of affect and this effect was overall stronger than that of blurring. However, the effect of blurring was stronger for unpleasant than for pleasant pictures, whereas the effect of coloring was stronger for pleasant than for unpleasant pictures.

In the analysis of the latencies, we found main and interaction effects of the control factor picture valence. Therefore, picture valence was included in the main analyses. The latency difference of affect and object recognition was positive in 38 of the 39 cases and averaged 469 ms. This difference was significant and was found for all combinations of task, coloring, and valence (coloring and valence were involved in a three-way interaction with the task factor). Therefore, Hypothesis 1 can be considered fully supported. Hypotheses 7a (object recognition precedes valence detection) and 7b (valence detection precedes affect) were also supported overall and in most cases also held true at the different levels of coloring, blurring, and picture valence (for exceptions, see Publication C, Franikowski & Reisenzein, 2022). In addition, blurring significantly delayed all latencies: object recognition and affect onset (H3) and valence detection (H9a). Consistent with predictions, coloring significantly delayed only affect (H5) and valence detection (H9b), but not object recognition.

Instead of computing only the between-subjects correlations between the focal latencies, the covariance structure was decomposed into a between-subjects and (based on the repeated measures due to the combinations of coloring, blurring, and repetitions) a within-subject part by estimating the multilevel correlation structure. That is, between-subjects and within-subject correlations were estimated. This was done separately for the pleasant and the unpleasant pictures. Whereas Hypothesis 2 (there is a positive correlation between the latencies of object recognition and affect) was supported at the within-subject level for both valences, at the between-subjects level it was only supported for pleasant pictures. Hypothesis 7a (there is a positive correlation between the latencies of object recognition and valence detection) was supported for the pleasant pictures at both the between- and within-level, but only for the unpleasant pictures at the between-level. Finally, Hypothesis 7b was supported for the pleasant but not for the unpleasant pictures (at both the between- and the within-level). To test mediation Hypotheses 4 and 6, multilevel path analyses were performed separately for the pleasant and unpleasant pictures. These analyses indicated that the delaying effect of blurring on affect latency was mediated, at least in part, by the delay in object recognition; this was found for both picture valences (H4). Moreover, the delaying effect of false-coloring on affect onset was found to be partly (for pleasant pictures) or completely (for unpleasant pictures) mediated by affect intensity (H6). Thus, both Hypotheses 4 and 6 were at least partially supported.

3.4.3 Discussion of Experiment Series C (Publication C)

Experiments C1 and C2 provided again support for the semantic primacy hypothesis of emotion generation by replicating the findings supporting Hypotheses 1 (there is a positive latency difference between affect and object recognition) and 2 (there is a positive correlation between these latencies; however, this support was only partial). Although picture valence interacted with the task factor in Experiment C2, picture valence only changed the effect size but not the direction of the task effect. All other control factors (task order, valence order, and blur group) had no effects. The hypotheses concerning the direct effects of blurring (H3) and false-coloring (H5) on the latencies of object recognition and affect were supported; the corresponding mediation hypotheses (H4: blurring delays affect via delaying object recognition; H6: false-coloring delays affect via reducing affect intensity) received partial support. Finally, the hypotheses concerning valence detection and its temporal relation to object recognition and affect (H7) were supported. The correlational hypothesis, that valence detection latency is positively correlated with the latencies of object recognition and affect (H8), and the experimental hypothesis, that both blurring and false-coloring delay valence detection latency (H9), were partially supported; it turned out that the effects depended on picture valence.

3.5 Meta-Analytic Integration of the Experiments

All eight experiments in the three Publications A, B, and C found that the temporal order of (H1) and the direction of the correlation between (H2) the object recognition and affect latencies were consistent with the semantic primacy hypothesis. In addition, the blurring manipulation (Experiments B2 and C2; H3 and H4) and the false-coloring manipulation (Experiment C2; H5 and H6) affected the object recognition and affect latencies, as well as the valence detection latencies (H9) in the predicted manner. Furthermore, valence detection preceded object recognition (H7a) and followed affect (H7b) and was positively correlated with both latencies (H8).

In agreement with the cross-method replication rationale, the latencies of object recognition and affect, as well as their latency differences and correlations, varied more or less as a function of the particular chronometric method used. To summarize the results obtained with the different methods for Hypotheses 1 and 2, I integrated them using meta-analytic methods. To maximize comparability, only data from the sharp and normally colored pictures from Experiments B2 and C2 were included. Depending on the chronometric method, the PSSs (Publication A), the 20% trimmed means (Publication B), and the medians (Publication C) served as latency estimates. Before integrating the effect sizes, the heterogeneity of effects due to the different families of chronometric methods was assessed within a random-effects model (Higgins et al., 2003); therefore, publication (chronometric method) was considered as a random factor. In the absence of substantial heterogeneity, the random-effects model could be simplified to a common-effect model. Analyses were performed in the *R* programming environment (R Core Team, 2022) using the *metamean* (analysis for H1) and *metacor* (analysis for H2) functions of the add-on package *meta* (Balduzzi et al., 2019).

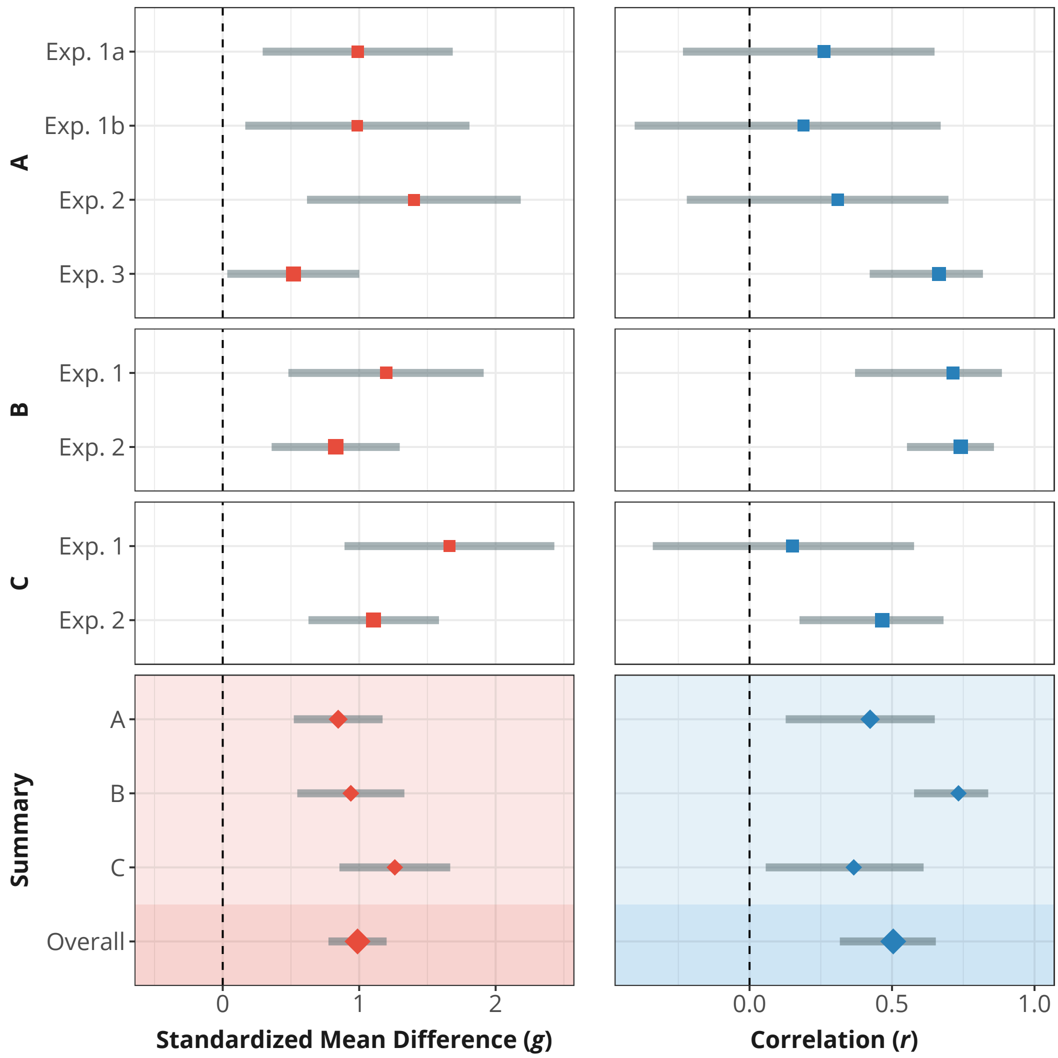
Regarding the differences between the latencies of affect and object recognition (H1), the temporal distances obtained in the speeded reaction time task cannot be compared in raw form to the temporal distance estimates obtained with the temporal judgment methods because reaction times and reaction time differences are much larger than the latencies and latency differences obtained with the temporal judgment methods. To overcome this difficulty, I used a standardized effect measure, Hedges' g, to combine the effects. To still obtain a cross-experiment estimate of the temporal distance between affect and object recognition, the temporal judgment estimates and the reaction time estimates were integrated separately. The correlations between the latencies (H2) were aggregated applying Fisher's r-to-z-transformation.

For Hypothesis 1, heterogeneity between the standardized effect sizes due to different chronometric methods was small and not significant $I^2 = 19\%$, $\tau^2 = 0.03$, $p = .278$; a test for differences between the methods (subgroups) also yielded a non-significant result, $\chi^2(2) = 1.63$, $p = .442$. Consequently, no random-effects structure was required for the estimation of the standardized latency differences. Therefore, the random-effects model was simplified to a common-effect model. For the aggregate model, the standardized mean difference indicated a large positive standardized effect, Hedges' $g = 0.99$, 95% CI [0.77, 1.20]. Because the estimate and its overall confidence interval are positive and represent a large effect, Hypothesis 1 was strongly supported. At the individual level, we found a positive latency difference between affect onset and object recognition in 176 of the 194 recruited participants. Integration of the raw mean differences yielded an aggregated latency difference of $M = 117$ ms, 95% CI [81, 153] ms for temporal judgments and $M = 378$ ms, 95% CI [149, 607] ms for speeded reaction times, using random-effects models, respectively.

For Hypothesis 2, the random-effects model revealed large, significant between-method heterogeneity, $I^2 = 52\%$, $\tau^2 = 0.05$, $p = .042$; the test for between-method differences was also significant, $\chi^2(2) = 8.02$, $p = .018$. Therefore, the random-effects structure was retained for estimating the cross-method effect size. The cross-method correlation coefficient was $r = .50$ with a consistently positive confidence interval, $r = .50$, 95% CI [.32, .65]. The largest, although insignificant, within-method heterogeneity was found in Publication A, $I^2 = 40\%$, $\tau^2 = 0.05$, $p = .172$, and reflects that the correlation coefficients obtained in the four experiments ranged from $r = .19$ to .66 (see Figure 3.2). However, based on the meta-analytic results, the correlation hypothesis was also strongly supported because the meta-analytic effect estimate was significantly positive and large.

Figure 3.2

Meta-Analytically Integrated Mean Differences and Correlation Coefficients



Note: The left panels show the standardized mean differences (Hedges' g) between the latencies of object recognition and affect and the overall difference (common-effect model); the right panels show the correlations between the two latencies and the overall correlation (random-effects model). The rows of the panels show the experiment-wise mean differences and correlations for the experiment series A, B, and C and, in the bottom row, the estimated summary effects per series (chronometric method) plus the overall (common-effect or random-effects) estimate across all experiments, nested within the different publications (and chronometric methods). The size of the squares and diamonds representing the point estimates of the effects reflects their relative weight in the meta-analytic aggregation. The horizontal bars indicate the 95% confidence intervals of the estimates.

Chapter 4

General Discussion

In the series of eight chronometric experiments reported in this dissertation, the semantic primacy hypothesis, which claims that object recognition is a necessary partial cause of affect, was unambiguously supported with three different methods of measuring perceptual latencies. All predictions derived from the semantic primacy hypothesis tested in our studies were supported. Specifically, all experiments confirmed the prediction that there is a positive mean difference between the latency of object recognition and affect (H1) and that the two latencies are positively correlated (H2). Even stronger support for semantic primacy was provided by the finding that an experimental manipulation that delayed object recognition, a moderate blurring of the affective pictures, also delayed affect (H3) and that this effect was partly mediated by delayed object recognition (H4).

In Experiments C1 and C2, we also investigated the relation of the latency of stimulus valence detection to the latencies of object recognition and affect onset. As predicted, the latency of valence detection was longer than the latency of object recognition and shorter than the latency of affect (H7) and correlated with both latencies (H8), indicating the interindividual consistency of these temporal order relations. Furthermore, similar to affect, valence detection was also delayed by blurring (H9a).

In addition to testing the semantic primacy hypothesis, Experiment C2 also tested the effects of false-coloring the affective pictures on the latencies of object recognition, affect, and valence detection. As predicted, we found that false-coloring delayed only the latency of affect but not that of object recognition (H5) and that this effect was mediated by reduced affect intensity (H6). This finding shows that affect intensity is another factor that influences the experienced onset of feelings. Furthermore, in parallel with the latency of affect, the latency of valence detection was delayed in the false-coloring condition (H9b). The correlation between the latencies of affect and valence detection and the parallel effects of blurring and false-coloring on these two variables provides some support for the decision of previous researchers to use the valence detection latency as a proxy for the onset of affect (Nummenmaa et al., 2010; Wells, 1925, 1929). Nevertheless, because valence judgments are not always based on experienced affect but can also be made by retrieving stored valence information from memory, the latency of valence judgments is a biased measure of affect onset. Seen from the perspective of cross-method and cross-measures replication, however, the mostly parallel (if, overall, weaker) effects obtained for valence judgments provide yet another piece of support for semantic primacy.

4.1 Comparison to Previous Findings

The results obtained for Hypotheses 1 and 2, specifically those of the speeded reaction time studies reported in Publication C, replicate the pattern of the findings of Nummenmaa et al. (2010). In particular, like these authors, we found a significant delay between object recognition and affect both overall and on the subject as well as on the stimulus level (H1) and a positive correlation between the latencies of object recognition and affect (H2; see especially Experiment 3 of Nummenmaa et al., 2010). The main difference between our findings and those of Nummenmaa et al. (2010) concerned the size of the cognition-affect lag, which was much smaller (40–80 ms) in the studies of Nummenmaa et al. (2010) than the difference found in our studies (meta-analytic integrated effect size in temporal judgment experiments: $M = 117$ ms; in the reaction time experiments: $M = 378$ ms). This difference in the findings was most likely due to methodological differences between the studies, in particular (1) the fact that Nummenmaa et al. (2010) measured the latency of valence detection rather than the latency of affect onset, and (2) the fact that they used a much shorter stimulus presentation time.

The importance of the first difference was empirically demonstrated in Experiments C1 and C2, which found that the latency of valence detection was significantly shorter than the latency of affect onset. As a consequence, the delay between object recognition and valence detection was also smaller than the latency difference between object recognition and affect.

However, even if this difference is considered, the cognition-affect delay reported by Nummenmaa et al. (2010) is still considerably smaller than that found in our studies C1 and C2. This brings us to the second factor: presentation time. Whereas Nummenmaa et al. (2010) presented the stimuli for only 30 ms, we presented our stimuli for at least 2,000 ms.[4] As a consequence of the extremely short presentation time, the participants in the studies of Nummenmaa et al. (2010) probably often felt uncertain about whether the stimulus was pleasant or not, or unpleasant or not; however, delaying the response to obtain more information about the stimulus made no sense because no further information was forthcoming. In contrast, in our studies, it made sense for the participants to wait until they were certain about the valence of the stimulus.

Because we used different chronometric methods to measure the latencies of object recognition and affect, the measurement methods contributed to the cross-study heterogeneity in latency estimates and latency differences. The most striking difference in this respect is that both latencies and latency differences were much larger in the speeded reaction time tasks (Experiments C1 and C2) than in the temporal judgment tasks. That the *reaction times* are longer than the perceptual latencies measured in the temporal judgment tasks is not surprising given that reaction times also contain motor decision and execution latencies. The larger *latency differences* obtained in the reaction time task, however, might be due to the fact that participants use a higher criterion for deciding on the presence of a signal in speeded reaction time tasks than in temporal judgment tasks (e.g., Cardoso-Leite & Gorea, 2010; Ejima & Ohtani, 1987; Miller & Schwarz, 2006). For a more extensive discussion, see also Publication C (Franikowski & Reisenzein, 2022).

[4]This comparatively long interval was chosen to imitate everyday situations in which an object can be viewed for a longer time than just 30 ms.

4.2 Possible Reasons for Deviating Response Patterns

The semantic primacy hypothesis was supported for the large majority of our participants but not for all. Specifically, for Hypothesis 1 (object recognition occurs before affect onset), we found that 12 of 81 participants in Publication A, 5 of 56 participants in Publication B and 1 of 57 participants in Publication C had a faster estimated latency for affect than for object recognition. One possible explanation of these deviating responses is that the semantic primacy hypothesis does not hold for all individuals but only in the (vast) majority, whereas a (small) minority conforms to the affective primacy hypothesis. That is, in these deviating participants, affect was evoked by a preconceptual representation of the stimuli.

To conclusively rule out that the deviating response patterns were due to affective primacy rather than being caused by measurement or other method problems, further investigation would be needed. A good approach would be to use a longitudinal design that measures object recognition and affect onset latencies repeatedly in the same individuals with different latency measurement methods. If response latencies suggesting affective primacy are consistently found with the different methods, an interindividual-differences version of the affective (and semantic) primacy hypothesis would be supported. In contrast, if there is no consistency across methods (affective primacy is only found with, say, temporal order judgments but not with speeded reaction times), then a measurement problem is the more likely explanation. Although we do not have the necessary longitudinal data, it is striking that rather different numbers of "deviant" participants were found in the different chronometric tasks: 15% in Publication A (temporal order and simultaneity judgments), 8% in Publication B (rotating spot task), and only 2% in Publication C (speeded reaction time task). Inasmuch as the participants of the different studies came from the same subject pool (predominantly undergraduate students, mostly of psychology, at the University of Greifswald), it is unlikely that these differences are due to sampling errors only.[5] This finding suggests that the temporal judgments, in particular the temporal order and simultaneity judgments, are more difficult than the speeded reaction time task, and that the deviating response patterns are due to participants who found a task too difficult.

Specifically, the reason for deviating latency differences (affect latency < object recognition latency) was probably that the affect task was too difficult for some participants. These participants may have responded randomly, or they could have resorted to a simplifying response strategy (see, e.g., Love et al., 2013). Supporting the first possibility, the data of eight of the excluded participants were very noisy and indicated near-random responding, especially in the case of temporal order and simultaneity judgments. Regarding the second possibility, participants who found the affect task too difficult could have replaced it with either appearance detection (one of the training tasks in most experiments), object recognition, or valence detection. If these participants resorted to appearance detection, their affect latencies would be smaller than the object recognition latencies, erroneously suggesting that the semantic primacy hypothesis is false. The data would also fail to confirm semantic primacy if these participants had judged when they recognized the object rather than when they felt the affect, as the mean difference in this case would have been (close to) 0 ms. Either case could explain the deviating response patterns. Finally, it

[5] A χ^2-test shows that these differences are significant, $\chi^2(1) = 5.24$, $p = .022$.

is possible that some participants judged the latency of detecting stimulus valence instead of the onset of the feeling; however, in this case, the mean latency difference to object recognition would still be positive and would be in line with the semantic primacy hypothesis, although the cognition-affect delay would be underestimated.

4.3 Implications of the Findings

The results of this dissertation research program are primarily relevant to the semantic versus affective primacy debate: They support semantic primacy, i.e., the hypothesis that object recognition is a necessary partial cause of emotion. In addition, however, the findings have implications for other questions of emotion research as well as for the chronometric study of emotional phenomena.

4.3.1 Affect Intensity as a Determinant of Affect Latency

Experiment C2 suggests that an additional factor that influences the latency of affect is the intensity of affect: False-coloring the pictures was found to reduce affect intensity as well as to delay affect onset, and a path analysis suggested that the latter effect was mostly mediated by reduced affect intensity.

However, the mediation findings should be interpreted with caution. First, replication is required. Second, full mediation of the effect of false-coloring on the latency of affect was only found for unpleasant stimuli, whereas for pleasant stimuli, the mediation was only partial, indicating that an additional process mediated the effect. One possibility is that the false-color manipulation altered the quality of the affect evoked by the pictures. For example, a child with a bluish face could have elicited pleasant feelings (because the picture showed a child) but also unpleasant feelings (because the bluish skin color may have suggested disease). In this case of mixed feelings, it may have required some additional time to focus attention on and judge the onset of the pleasant feeling (as required by the instruction used in Experiment C2). Alternatively, the mechanisms that generate affect may need more time to compute the affect for false-colored objects because the information about the category (e.g., child) and the information about unusual object features (e.g., bluish face) are difficult to integrate. To clarify these issues, a cleaner manipulation of affect intensity should be used. One promising manipulation would be picture size (e.g., De Cesarei & Codispoti, 2008; Junge & Reisenzein, 2013).

4.3.2 Processes Underlying Valence Detection

A secondary goal of this research program was to clarify the processes underlying valence judgments; among other reasons because these are frequently used as proxies for affective experiences (see, e.g., Itkes et al., 2017; Nummenmaa et al., 2010). We hypothesized that valence detection is based partly on the feelings evoked by the pictures and partly on the retrieval of valence information stored in object schemata. Specifically, if the stimuli are well-known, the valence judgments are based on memory and their latency is close to that of object recognition. If the stimuli are novel, in contrast, the valence judgments are based on feelings and their latency is close to that of affect.

To maximize the chances of memory-based valence judgments, we presented the valence judgment task as the last task in the speeded reaction time experiments (C1 and C2). The results are in line with the assumption that the valence judgments were indeed in part memory-based; however, a more stringent test of this hypothesis is desirable. Extending the logic of Experiments C1 and C2, such a test could be made by comparing the latency of valence judgments for stimulus sets consisting of different proportions of novel vs. well-known stimuli, as well as under different conditions (instructions emphasizing vs. deemphasizing speed; long vs. short stimulus presentation). In theory, higher proportions of known stimuli, "fast response" instructions and short stimulus durations should shift the latency of valence detection closer to the latency of object recognition.

Another potential criticism is that the valence detection task was a one-alternative forced choice (1AFC) reaction time task, whereas a simple reaction time task was used in the object recognition and affect tasks. It could be argued that the shorter latency of valence detection, compared to the latency of affect, was in fact caused by this difference between the tasks. For example, the participants could have relied exclusively on affect when making the valence judgments but may have used a lower decision criterion in the valence judgment task (analogous to the lower criterion in temporal judgment compared to speeded reaction time tasks; e.g., Miller & Schwarz, 2006), resulting in shorter latencies. However, it is unclear why different decision criteria should have been used in two different reaction time tasks. Furthermore, other factors being constant, forced-choice reaction time tasks take more time than simple reaction time tasks (e.g., Miller & Low, 2001).

In summary, the nature and processes underlying valence detection require systematic future investigation. In these future studies, it may be advisable to measure the latencies of object recognition, affect, and valence detection using the same kind of reaction time task (i.e., simple, one-alternative forced choice, or two-alternative-forced choice as in Nummenmaa et al., 2010).

4.3.3 Bodily Feedback Theories of Emotional Experience

The results of this dissertation project also have implications for theories of bodily feedback in emotional experience. According to a tradition of emotion theorizing going back to James (1890), bodily feedback is necessary for emotional experience, or at least contributes significantly to its quality and / or intensity (Damasio, 1994; James, 1890; Laird, 2007; Schachter, 1964). However, critics of bodily feedback theories have raised a variety of objections to these assumptions. One of these objections, apparently first raised by Stumpf (1899) and later reiterated more forcefully by Cannon (1927), is that the latency of physiological responses evoked by affective stimuli is too long for them to be the basis of emotional experience. In defense, proponents of the bodily feedback theory have argued that at least some physiological responses evoked by emotional stimuli are shorter than those investigated by Cannon (1927; e.g., changes in heart rate can be as short as 1,000 ms, see Bradley et al., 2001) and that facial (corrugator and zygomatic) muscle responses to affective pictures have been reported to occur as early as 500 ms after stimulus onset (Dimberg et al., 1998).

Concerning physiological reactions, even a latency as short as 1,000 ms is too long for bodily feedback to fit comfortably into the time window allowed by the reaction time estimates of affect onset provided by Nakashima (1909a, 1909b), Nummenmaa et al. (2010), and Wells (1925), and studies C1 and C2 of this dissertation project; in particular, if one takes into account that (a) the affect latency estimates obtained in the reaction time studies are definitely too long, because they also include motor decision and execution time and (b) the time window allowed by these estimates must not only cover the elicitation of the physiological reaction but also the time it takes to report this reaction back to the brain (Cannon, 1927).

However, facial reactions to emotional stimuli could meet the temporal constraints imposed by the latency of affect estimated by the reaction time studies because these estimates are fairly imprecise. Our temporal judgment studies, however, provide more precise latency estimates. Although these estimates are relative (to the onset of the probe) rather than absolute as reaction time measurements are, absolute estimates can be obtained if the absolute occurrence time of but one of the measured latencies is known. As it turns out, relatively precise estimates of the absolute latency of object recognition are available from studies that used eye saccades as the response mode because these responses can be executed extremely quickly. Using this method, Kirchner and Thorpe (2006) estimated that object recognition occurs about 230 ms after stimulus presentation, while Nummenmaa et al. (2010) estimated a latency of 290 ms (for a recent review, see Förster et al., 2020). Subtracting an estimate of 30 ms for the motor execution time, i.e., performing the saccade (Kirchner & Thorpe, 2006; Schiller & Kendall, 2004), the latency of recognizing natural objects is between 200 and 260 ms. This time range agrees well with the object recognition latency estimated in ERP (event-related potential) studies (Fabre-Thorpe, 2011; Johnson & Olshausen, 2003). Adding the delay of about 117 ms (derived from the meta-analytic integration of the temporal judgment experiments of Publications A and B), the latency of the emotional experience ranges from 317 (200 + 117) to 377 (260 + 117) ms after picture onset.

Interestingly, the lower limit of this range is consistent with the latency of the potentiation of the startle response by aversive pictures (Globisch et al., 1999) and the latency of the affect modulation of the LPP (late positive potential) component of the ERP (Hajcak & Foti, 2020). However, even the upper limit of the affect latency range (377 ms) would be too short for facial feedback (and even more, physiological feedback) to be the basis of affect.

Proponents of feedback theories of emotion might object that their hypotheses may still hold (and perhaps were always meant to hold) for basic emotions (Ekman, 1992). However, our findings suggest that this objection is invalid at least for the case of disgust, as in most of our studies, the unpleasant picture-induced feeling was actually disgust.

4.4 Limitations and Directions for Future Research

In this dissertation project, the different latency measurement methods were varied between participants, i.e., each participant was tested with only one measurement method. An important task for follow-up research is to test to what degree the latency estimates obtained with the different chronometric methods are intraindividually consistent (among other reasons to clarify the reasons for deviating response patterns). To answer this question, it is necessary to time a participant's mental events (object recognition, affect, and valence detection) using multiple chronometric methods. Such a study would of course require longer or distributed experimental sessions. Longer sessions can cause tiredness and a loss of attention in participants, whereas multiple session designs are subject to state effects (e.g., concentration and wakefulness varying across the sessions) and potential dropouts. Because of the large trial number they require, temporal order and simultaneity judgments are less suitable for this cross-method approach than the rotating spot method and speeded reaction time tasks.

The studies reported in this project address only the broad categories of pleasant and unpleasant emotional experiences. A logical extension of the latency measurement methods (in particular the temporal judgment methods) would be to study more specific emotional experiences. Promising candidates apart from disgust (which, it can be argued, was already considered implicitly in our studies) might be surprise, amusement, and curiosity, because these feelings can be easily and repeatedly induced in the laboratory (e.g., Brod et al., 2018; Fiacconi & Owen, 2015; Reisenzein, 2000). In addition, the chronometric methods used in our studies can be used to time yet other emotion-related events and to compare their latencies. This can provide for novel tests of several other hypotheses of emotion psychology. Chronometric methods could be used, for example, to decide between the conflicting hypotheses that emotional experiences *are* felt action tendencies (e.g., Arnold, 1960; Frijda, 1986) versus *cause* them (e.g., Mellers, 2000; Weiner, 1995).

In our experiments, the stimuli were limited to pictorial affective stimuli. A natural extension of the present research, therefore, would be to study affective stimuli from other sensory modalities. Nakashima (1909a) already used tones to time object recognition (i.e., classification of tones as low, medium or high) and tone-evoked affects. His results again supported semantic primacy; however, it would be important to replicate these findings and also include more complex affective sounds from everyday life.

In addition to the experimental manipulations of the affective pictures realized in Experiments B2 and C2, blurring and false-coloring, yet other picture modifications could be used in future studies to test the stability of our findings and to further clarify the factors that influence affect latency. For example, as already mentioned, to obtain a cleaner manipulation of affect intensity than that provided by false-coloring, the size of the affective pictures could be manipulated. Picture size has a reliable effect on affect intensity (Junge & Reisenzein, 2013) whereas object recognition should not be hampered as long as the pictures are not too small. In contrast to false-coloring, the quality of the affect should not be influenced either. In experiments that test the semantic primacy hypothesis with affective sounds instead of affective pictures, object recognition could be delayed, for example, by overlaying the sound with white noise. Corresponding to the picture size manipulation in the visual domain, for example, the volume of the sounds could be manipulated. This may be expected to reduce affect intensity without delaying object recognition (as long as the volume is not too low).

In the studies reported in this dissertation, we only investigated the *onset* of the perceptual latencies of object recognition and affect. Going beyond this, one may also be interested in the *dynamics* of affective feelings, i.e., not only their onset but also their progression (e.g., the time until they reach a maximum and the time until offset). The time course of emotional intensity has already been assessed using different measurement approaches (Oceja & Carrera, 2009; Sonnemans & Frijda, 1994; Verduyn et al., 2009). These approaches could be used to check whether our participants really reported the onset of affect, or instead another intensity parameter of the course of affect across time, such as the maximum of affect intensity or a point beyond the onset but before the maximum. This measurement approach would also be valuable in studying the course of affect for dynamic (e.g., videos, sounds) rather than static (e.g., pictures) stimuli.

Finally, in the studies reported here, we timed the onset of emotional experience using introspection-based measurement methods. Because we were interested in subjective experience and introspection provides the most direct access to emotional experience, this was the natural choice. However, these methods, in particular the temporal judgment methods, could be profitably combined in future studies with objective data such as physiological reactions (e.g., heart rate, skin conductance), facial responses, or evoked brain potentials (e.g., the LPP). Because of its economy, the rotating spot task may be particularly well suited for the parallel recording of physiological and expressive reactions. These studies could shed further light on the semantic primacy hypothesis, as well as more broadly on the temporal and causal relations between the mental and bodily processes involved in emotions.

4.5 Conclusion

The dissertation project began as a revival of introspectionist psychology, taking up and extending the research program of Taizo Nakashima (1909a, 1909b), which suffered from methodological limitations and focused only on examining the temporal order of object recognition and affect onset. To address these shortcomings, the three series of studies presented in this dissertation (1) used four different, modern chronometric methods, (2) tested not only the temporal order of object recognition and affect latency but also its interindividual consistency, and (3) extended the latency measurement paradigm by including an experimental test of the semantic primacy hypothesis. In addition, modern statistical methods were used to analyze the data and sufficiently large sample sizes were recruited to permit testing the central hypotheses with high power (see Figure 2.1). Finally, the findings of the eight experiments were meta-analytically integrated in this introduction. The findings consistently support the semantic primacy hypothesis across different latency measurement methods and they also shed light on the processes underlying valence judgments. It is hoped that this research program will stimulate future chronometric research in emotion psychology; several possible routes for such research were proposed.

References

Arnold, M. B. (1960). *Emotion and personality*. Columbia University Press.

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health*, *22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Barrett, L. F., & Bliss–Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, *41*, 167–218. https://doi.org/10.1016/S0065-2601(08)00404-8

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–298. https://doi.org/10.1037/1528-3542.1.3.276

Brod, G., Hasselhorn, M., & Bunge, S. A. (2018). When generating a prediction boosts learning: The element of surprise. *Learning and Instruction*, *55*, 22–31. https://doi.org/10.1016/j.learninstruc.2018.01.013

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology*, *39*(1/4), 106. https://doi.org/10.2307/1415404

Cardoso-Leite, P., & Gorea, A. (2010). On the perceptual/motor dissociation: A review of concepts, theory, experimental paradigms and data interpretations. *Seeing and Perceiving*, *23*(2), 89–151. https://doi.org/10.1163/187847510X503588

Carlson, T. A., Hogendoorn, H., & Verstraten, F. A. J. (2006). The speed of visual attention: What time is it? *Journal of Vision*, *6*(12), 1406–1411. https://doi.org/10.1167/6.12.6

Clore, G. L., & Byrne, D. (1974). A reinforcement-affect model of attraction. In T. L. Huston (Ed.), *Foundations of interpersonal attraction* (pp. 143–170). Academic Press. https://doi.org/10.1016/B978-0-12-362950-0.50013-6

Codispoti, M., Ferrari, V., & Bradley, M. M. (2006). Repetitive picture processing: Autonomic and cortical correlates. *Brain Research*, *1068*(1), 213–220. https://doi.org/10.1016/j.brainres.2005.11.009

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology*, *46*(4), 731–738. https://doi.org/10.1111/j.1469-8986.2009.00804.x

Crane, T. (Ed.). (2009). *The contents of experience: Essays on perception*. Cambridge University Press. https://doi.org/10.1017/CBO9780511554582

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95–111. https://doi.org/10.1016/S0301-0511(99)00044-7

Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain.* Avon.

Danquah, A. N., Farrell, M. J., & O'Boyle, D. J. (2008). Biases in the subjective timing of perceptual events: Libet et al. (1983) revisited. *Consciousness and Cognition*, *17*(3), 616–627. https://doi.org/10.1016/j.concog.2007.09.005

De Cesarei, A., & Codispoti, M. (2008). Fuzzy picture processing: Effects of size reduction and blurring on emotional processing. *Emotion*, *8*(3), 352–363. https://doi.org/10.1037/1528-3542.8.3.352

De Cesarei, A., & Codispoti, M. (2010). Effects of picture size reduction and blurring on emotional engagement. *PloS one*, *5*(10), e13399. https://doi.org/10.1371/journal.pone.0013399

Diederich, A., & Colonius, H. (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order. *Psychological Review*, *122*(2), 232–241. https://doi.org/10.1037/a0038696

Dimberg, U., Hansson, G. O., & Thunberg, M. (1998). Fear of snakes and facial reactions: A case of rapid emotional responding. *Scandinavian Journal of Psychology*, *39*(2), 75–80. https://doi.org/10.1111/1467-9450.00059

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89. https://doi.org/10.1111/1467-9280.00221

Ejima, Y., & Ohtani, Y. (1987). Simple reaction time to sinusoidal grating and perceptual integration time: Contributions of perceptual and response processes. *Vision Research*, *27*(2), 269–276. https://doi.org/10.1016/0042-6989(87)90189-1

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4), 169–200. https://doi.org/10.1080/02699939208411068

Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology*, *2*, 243. https://doi.org/10.3389/fpsyg.2011.00243

Fazio, R. H. (2007). Attitudes as object-evaluation Associations of varying strength. *Social Cognition*, *25*(5), 603–637. https://doi.org/10.1521/soco.2007.25.5.603

Ferrari, V., Mastria, S., & Codispoti, M. (2020). The interplay between attention and long-term memory in affective habituation. *Psychophysiology*, *57*(6), e13572. https://doi.org/10.1111/psyp.13572

Fiacconi, C. M., & Owen, A. M. (2015). Using psychophysiological measures to examine the temporal profile of verbal humor elicitation. *PloS one*, *10*(8), e0135902. https://doi.org/10.1371/journal.pone.0135902

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, *117*(1), 39–66. https://doi.org/10.1037/0033-2909.117.1.39

Förster, J., Koivisto, M., & Revonsuo, A. (2020). ERP and MEG correlates of visual consciousness: The second decade. *Consciousness and Cognition*, *80*, 102917. https://doi.org/10.1016/j.concog.2020.102917

Franikowski, P., Kriegeskorte, L.-S., & Reisenzein, R. (2021). Perceptual latencies of object recognition and affect measured with the rotating spot method: Chronometric evidence for semantic primacy. *Emotion*, *21*(8), 1744–1759. https://doi.org/10.1037/emo0000991

Franikowski, P., & Reisenzein, R. (2022). On the latency of object recognition and affect: Evidence from speeded reaction time tasks. *Emotion.* Advance online publication. https://doi.org/10.1037/emo0001092

Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
García-Pérez, M. A., & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. *Psychonomic Bulletin & Review*, *19*(5), 820–846. https://doi.org/10.3758/s13423-012-0278-y
Geiger, M. (1903). Neue Complicationsversuche [New complication experiments]. *Philosophische Studien*, *18*, 347–436.
Globisch, J., Hamm, A. O., Esteves, F., & Öhman, A. (1999). Fear appears fast: Temporal course of startle reflex potentiation in animal fearful subjects. *Psychophysiology*, *36*(1), 66–75. https://doi.org/10.1017/S0048577299970634
Haering, C., & Kiesel, A. (2012). Mine is earlier than yours: Causal beliefs influence the perceived time of action effects. *Frontiers in Psychology*, *3*, 393. https://doi.org/10.3389/fpsyg.2012.00393
Hajcak, G., & Foti, D. (2020). Significance? ... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology*, *57*(7), e13570. https://doi.org/10.1111/psyp.13570
Hedger, N., Adams, W. J., & Garner, M. (2015). Autonomic arousal and attentional orienting to visual threat are predicted by awareness. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 798–806. https://doi.org/10.1037/xhp0000051
Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557
Hume, D. (2009). *A treatise of human nature*. Oxford University Press. (Original work published 1749)
Itkes, O., Kimchi, R., Haj-Ali, H., Shapiro, A., & Kron, A. (2017). Dissociating affective and semantic valence. *Journal of Experimental Psychology: General*, *146*(7), 924–942. https://doi.org/10.1037/xge0000291
James, W. (1890). *The principles of psychology* (Vol. 2). Holt.
Jaśkowski, P. (2014). What determines simultaneity and order perception? In V. Artstila & D. Lloyd (Eds.), *Subjective time: The philosophy, psychology, and neuroscience of temporality* (pp. 379–407). MIT Press.
Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision*, *3*(7), 499–512. https://doi.org/10.1167/3.7.4
Junge, M., & Reisenzein, R. (2013). Indirect scaling methods for testing quantitative emotion theories. *Cognition & Emotion*, *27*(7), 1247–1275. https://doi.org/10.1080/02699931.2013.782267
Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776. https://doi.org/10.1016/j.visres.2005.10.002
Kosovicheva, A., & Bex, P. J. (2020). What color was it? A psychophysical paradigm for tracking subjective progress in continuous tasks. *Perception*, *49*(1), 21–38. https://doi.org/10.1177/0301006619886247
Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General*, *144*(2), 339–365. https://doi.org/10.1037/xge0000040

Laird, J. D. (2007). *Feelings*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195098891.001.0001

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, *37*(9), 1019–1024. https://doi.org/10.1037/0003-066X.37.9.1019

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.

Lazarus, R. S. (1995). Vexing research problems inherent in cognitive-mediational theories of emotion - and some solutions. *Psychological Inquiry*, *6*(3), 183–196. https://doi.org/10.1207/s15327965pli0603_1

LeDoux, J. E. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon & Schuster.

Leys, R. (2017). *The ascent of affect*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226488738.001.0001

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain: A Journal of Neurology*, *106*(3), 623–642. https://doi.org/10.1093/brain/106.3.623

Linares, D., & Holcombe, A. O. (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent. *i-Perception*, *5*(6), 559–571. https://doi.org/10.1068/i0675

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PloS one*, *8*(1), e54798. https://doi.org/10.1371/journal.pone.0054798

Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Clarendon Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. The Freeman.

Marsella, S., Gratch, J., & Petta, P. (2010). Computational models of emotion. In K. R. Scherer, T. Bänziger, & E. Roesch (Eds.), *A blueprint for affective computing: A sourcebook and manual* (pp. 21–46). Oxford University Press.

Meinong, A. (1984). *Psychologisch-ethische Untersuchungen zur Werttheorie [Psychological-ethical investigations in value theory]*. Leuschner & Lubensky.

Mellers, B. A. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin*, *126*(6), 910–924. https://doi.org/10.1037/0033-2909.126.6.910

Micucci, A., Ferrari, V., De Cesarei, A., & Codispoti, M. (2020). Contextual modulation of emotional distraction: Attentional capture and motivational significance. *Journal of Cognitive Neuroscience*, *32*(4), 621–633. https://doi.org/10.1162/jocn_a_01505

Miller, J. O., & Low, K. (2001). Motor processes in simple, go/no-go, and choice reaction time tasks: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(2), 266–289. https://doi.org/10.1037/0096-1523.27.2.266

Miller, J. O., & Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: A diffusion model approach. *Journal of Exper-*

imental Psychology: Human Perception and Performance, *32*(2), 394–412. https://doi.org/10.1037/0096-1523.32.2.394

Miller, J. O., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review*, *20*(5), 819–858. https://doi.org/10.3758/s13423-013-0404-5

Mitschka, J. J. (2021). *Die Latenz von Gefühlen, gemessen mit der "Direkten Reaktionsmethode" von Nakashima [The latency of feelings measured with Nakashima's "direct rection method"]* [Unpublished bachelor's thesis]. Department of Psychology, University of Greifswald.

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*(6684), 467–470. https://doi.org/10.1038/30976

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*(5), 723–739. https://doi.org/10.1037/0022-3514.64.5.723

Nakashima, T. (1909a). Contributions to the study of the affective processes. *The American Journal of Psychology*, *20*(2), 157. https://doi.org/10.2307/1413295

Nakashima, T. (1909b). Time-relations of the affective processes. *Psychological Review*, *16*(5), 303–339. https://doi.org/10.1037/h0075294

Neumann, O., & Niepel, M. (2004). Timing of perception and perception of time. In C. Kaernbach, E. Schröger, & H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 245–269). Erlbaum.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General*, *139*(2), 222–246. https://doi.org/10.1037/a0018858

Oceja, L., & Carrera, P. (2009). Beyond a single pattern of mixed emotional experience. *European Journal of Psychological Assessment*, *25*(1), 58–67. https://doi.org/10.1027/1015-5759.25.1.58

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522. https://doi.org/10.1037/0033-295X.108.3.483

Öhman, A., & Soares, J. J. F. (1994). "Unconscious anxiety": Phobic responses to masked stimuli. *Journal of Abnormal Psychology*, *103*(2), 231–240. https://doi.org/10.1037/0021-843X.103.2.231

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press. https://doi.org/10.1017/CBO9780511571299

Peira, N., Golkar, A., Öhman, A., Anders, S., & Wiens, S. (2012). Emotional responses in spider fear are closely related to picture awareness. *Cognition & Emotion*, *26*(2), 252–260. https://doi.org/10.1080/02699931.2011.579087

Pessoa, L., Japee, S., & Ungerleider, L. G. (2005). Visual awareness and the detection of fearful faces. *Emotion*, *5*(2), 243–247. https://doi.org/10.1037/1528-3542.5.2.243

Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition*, *16*(2), 241–254. https://doi.org/10.1016/j.concog.2006.09.002

Popper, K. R. (1959). *The logic of scientific discovery*. Basic Books.

Posner, M. I. (1986). *Chronometric explorations of mind.* Oxford University Press.

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Reisenzein, R. (2000). The subjective experience of surprise. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 262–279). Psychology Press.

Reisenzein, R. (2009). Emotions as metarepresentational states of mind: Naturalizing the belief–desire theory of emotion. *Cognitive Systems Research*, *10*(1), 6–20. https://doi.org/10.1016/j.cogsys.2008.03.001

Reisenzein, R., & Franikowski, P. (2022). On the latency of object recognition and affect: Evidence from temporal order and simultaneity judgments. *Journal of Experimental Psychology: General.* Advance online publication. https://doi.org/10.1037/xge0001244

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Scarantino, A. (2010). Insights and blindspots of the cognitivist theory of emotions. *The British Journal for the Philosophy of Science*, *61*(4), 729–768. https://doi.org/10.1093/bjps/axq011

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 48–81). Academic Press.

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1535), 3459–3474. https://doi.org/10.1098/rstb.2009.0141

Schiller, P. H., & Kendall, J. (2004). Temporal factors in target selection with saccadic eye movements. *Experimental Brain Research*, *154*(2), 154–159. https://doi.org/10.1007/s00221-003-1653-8

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, *47*(4), 333–366. https://doi.org/10.1016/S0010-0285(03)00035-5

Schupp, H. T., Stockburger, J., Schmälzle, R., Bublatzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: Brain potentials and stimulus identification. *Neuroreport*, *19*(2), 167–171. https://doi.org/10.1097/WNR.0b013e3282f4aa42

Seifried, T., Ulrich, R., Bausenhart, K. M., Rolke, B., & Osman, A. (2010). Temporal preparation decreases perceptual latency: Evidence from a clock paradigm. *Quarterly Journal of Experimental Psychology*, *63*(12), 2432–2451. https://doi.org/10.1080/17470218.2010.485354

Sonnemans, J., & Frijda, N. H. (1994). The structure of subjective emotional intensity. *Cognition & Emotion*, *8*(4), 329–350. https://doi.org/10.1080/02699939408408945

Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*(5), 543–545. https://doi.org/10.1038/nn.2112

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799–832. https://doi.org/10.1037/0096-3445.130.4.799

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and perfomance IV* (pp. 629–685). Academic Press.

Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic processing precedes affect retrieval: The neurological case for cognitive primacy in visual processing. *Review of General Psychology*, *10*(1), 41–55. https://doi.org/10.1037/1089-2680.10.1.41

Stumpf, C. (1899). Über den Begriff der Gemüthsbewegung [On the concept of emotion]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *21*, 47–99.

van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics*, *70*(6), 955–968. https://doi.org/10.3758/PP.70.6.955

Verbaarschot, C., Haselager, P., & Farquhar, J. (2016). Detecting traces of consciousness in the process of intending to act. *Experimental Brain Research*, *234*(7), 1945–1956. https://doi.org/10.1007/s00221-016-4600-1

Verduyn, P., van Mechelen, I., Tuerlinckx, F., Meers, K., & van Coillie, H. (2009). Intensity profiles of emotional experience over time. *Cognition & Emotion*, *23*(7), 1427–1443. https://doi.org/10.1080/02699930902949031

Watson, J. B. (1919). *Psychology from the standpoint of a behaviorist*. Lippincott. https://doi.org/10.1037/10016-000

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. Guilford.

Weiß, K., Hilkenmeier, F., & Scharlau, I. (2013). Attention and the speed of information processing: Posterior entry for unattended stimuli instead of prior entry for attended stimuli. *PloS one*, *8*(1), e54257. https://doi.org/10.1371/journal.pone.0054257

Wells, F. L. (1925). Reactions to visual stimuli in affective settings. *Journal of Experimental Psychology*, *8*(1), 64–76. https://doi.org/10.1037/h0066415

Wells, F. L. (1929). Reaction-times to affects accompanying smell stimuli. *The American Journal of Psychology*, *41*(1), 83. https://doi.org/10.2307/1415110

Wilcox, R. R. (2016). *Introduction to robust estimation and hypothesis testing* (4th ed.). Elsevier.

Wundt, W. (1897). *Grundzüge der physiologischen Psychologie [Outlines of physiological psychology]* (3rd ed., Vol. 2). Engelmann.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Publications in Peer-Reviewed Journals

Publication A

Reisenzein, R., & Franikowski, P. (2022). On the latency of object recognition and affect: Evidence from temporal order and simultaneity judgments. *Journal of Experimental Psychology: General.* Advance online publication. https://doi.org/10.1037/xge0001244

Publication B

Franikowski, P., Kriegeskorte, L.-S., & Reisenzein, R. (2021). Perceptual latencies of object recognition and affect measured with the rotating spot method: Chronometric evidence for semantic primacy. *Emotion*, *21*(8), 1744–1759. https://doi.org/10.1037/emo0000991

Publication C

Franikowski, P., & Reisenzein, R. (2022). On the latency of object recognition and affect: Evidence from speeded reaction time tasks. *Emotion.* Advance online publication. https://doi.org/10.1037/emo0001092

Appendix A

Publication A

On the latency of object recognition and affect: Evidence from temporal order and simultaneity judgments

Reisenzein, R.*, & Franikowski, P.* (2022).
Journal of Experimental Psychology: General. Advance online publication.
https://doi.org/10.1037/xge0001244
* Shared first authorship.

The final article is available, upon publication, at: https://doi.org/10.1037/xge0001244

On the Latency of Object Recognition and Affect: Evidence from Temporal Order and Simultaneity Judgments

Rainer Reisenzein* and Philipp Franikowski*
University of Greifswald

Abstract

According to the semantic primacy hypothesis of emotion generation, stimuli must be semantically categorized to evoke emotions. This hypothesis was tested for the subjective component of emotions in four chronometric experiments in which the conscious recognition of emotion-eliciting objects and the onset of affect was timed using temporal order judgments (TOJs, Exp. 1a, 1b and 3) and simultaneity judgments (SJs, Exp. 2). Participants viewed pictures that elicited pleasant or unpleasant feelings. At varying intervals before and after picture onset, a visual probe stimulus was presented. In separate blocks of trials, the participants judged when they recognized the object shown in the picture and noticed the feeling evoked by the picture: Before or after the probe (TOJs), or simultaneously/not simultaneously with the probe (SJs). Psychometric functions were fitted to the data of the individual participants to determine the point of subjective simultaneity (PSS) of the target events and the probe. In both tasks, the mean PSS of affect occurred significantly later than the PSS of object recognition, with an averaged delay of about 120 ms. A positive lag between object recognition and affect was found for nearly all participants and pictures. In addition, the latencies of object recognition and affect were positively correlated. The found temporal order of object recognition and affect is consistent with the findings of previous studies using speeded reaction time tasks. Implications of the findings for the cognition-emotion debate and for bodily feedback theories of emotional experience are discussed.

Keywords: latency of affect; latency of object recognition; semantic and affective primacy; temporal order judgments; simultaneity judgments

Since the pioneering studies of Exner (1875) and Wundt (1897), measurements of the subjective occurrence times of perceptions and other kinds of conscious experiences have become a much-used tool for investigating the temporal dimension of consciousness (see e.g., Jaśkowski, 2014; Kostaki & Vatakis, 2018; García-Pérez & Alcalá-Quintana, 2012; Tünnermann et al., 2017). The basic idea of these measurement methods is to ask participants to relate the occurrence of a mental event of interest (e.g., the perception of a tone) to that of a contemporaneously

---

Author note

Rainer Reisenzein, https://orcid.org/0000-0002-6091-108X, Philipp Franikowski, https://orcid.org/0000-0002-2109-7388, Institute of Psychology, University of Greifswald
Correspondence concerning this article should be addressed to Rainer Reisenzein, Institute of Psychology, University of Greifswald, Franz-Mehring-Str. 47, 17489 Greifswald, Germany. E-mail: rainer.reisenzein@uni-greifswald.de
All analyzed data and R scripts are available on https://osf.io/eugc5/.
This study was not preregistered.
* The two authors contributed equally to this work.

occurring comparison event (e.g., the perception of a light). In one variant of the temporal judgment method, pioneered by Wundt (1897) and Geiger (1903), the comparison events are the perceptions produced by a continuous stimulus stream, such as a swinging pendulum or a fast-moving clock hand, and the participant's task is to report the position of the comparison stimulus when the target event occurs. This method continues to be used successfully to the present day (e.g., Libet et al., 1983; Pockett & Miller, 2007; Weiß et al., 2013). More often, however, researchers have followed Exner (1875) in using simpler temporal judgment tasks that presumably are easier for the participants to make. The most frequently used tasks of this kind are temporal order judgments (TOJs) and simultaneity judgments (SJs; for reviews, see e.g., García-Pérez & Alcalá-Quintana, 2012; Jaśkowski, 2014; Kostaki & Vatakis, 2018; Neumann & Niepel, 2004; Sternberg & Knoll, 1973).

The TOJ and SJ tasks are adaptations of psychophysical threshold detection methods – typically, the method of constant stimuli – to the perception of occurrence times. In both tasks, a stimulus that reliably evokes a mental state of interest is preceded and followed, at different short intervals, by another stimulus, and the participant is asked to make a judgment about the temporal relation between the experiences evoked by the two stimuli.[1] We will henceforth call the stimulus that evokes the mental event of interest the *target* stimulus and the other stimulus the *probe* stimulus, to emphasize that our interest is on the target stimulus, or more precisely, on certain mental events evoked by it (object recognition and the experience of affect), whereas the probe is of interest only as a necessary standard of comparison. In the TOJ task, the participants are instructed to report which mental event they became aware of first, or whether they became aware of the target event before or after the probe (cf. Sternberg & Knoll, 1973); in the SJ task, they are asked to report whether or not the target and probe event appeared to be simultaneous (see e.g., García-Pérez & Alcalá-Quintana, 2012; Sternberg & Knoll, 1973; van Eijk et al., 2008). The data generated by the two tasks are summarized in proportions of "target first" (or "target before probe") judgments in the TOJ task, and proportions of "simultaneous" judgments in the SJ task, one for each of a set of experimentally manipulated time intervals (SOAs, stimulus onset asynchronies) between target and probe stimulus. These data are then typically used to estimate the (positive or negative) delay between the target and probe stimulus at which the mental events caused by them are experienced as simultaneous; this is called the *point of subjective simultaneity* (PSS). For TOJs, the PSS is the delay of the probe stimulus at which the probability of the judgments "target event before probe" and "target event after probe" is equal (.50); for SJs, the PSS is the temporal delay of the probe where the probability of the "simultaneous" judgment peaks. If the onset time of the target stimulus is set to 0, the SOA for the PSS directly reflects the time difference between the judged onset of the target event and that evoked by the probe. Hence, TOJ and SJ tasks can be used to measure the relative occurrence time of a target mental event (i.e., its occurrence time relative to the probe event; Schneider & Bavelier, 2003).

Based on this reasoning, TOJ and SJ tasks have been used to investigate numerous

[1] Today's temporal judgment researchers typically speak of temporal relations between *perceived stimuli* (e.g., the onset of a light and a tone) rather than between *perceptions of stimuli.* Nonetheless, what is compared by the subject are the occurrence times of conscious experiences (e.g., Wundt, 1897; Sternberg & Knoll, 1973). Furthermore, the "objectifying" manner of speech cannot be used for mental events such as affective experiences (our studies) or action intentions (Libet et al. 1983).

research questions, mostly in the psychology of perception, for which relative onset times are relevant (see e.g., Diederich & Colonius, 2015; Haering & Kiesel, 2012; Jaśkowski, 2014; Sternberg & Knoll,1973).

The Temporal Order of Object Recognition and Affect: Semantic or Affective Primacy?

In the studies reported in this article, we put temporal order and simultaneity judgments to a new task: We used them to estimate the relative onset times of (1) pleasant or unpleasant affect (feelings) evoked by emotional pictures, and (2) the conscious recognition of the objects shown in the pictures. Information about the timing, in particular the temporal order, of the mental events involved in emotional processes is relevant to many questions of emotion psychology (e.g., Globisch et al., 1999; Scherer, 2009; Siemer & Reisenzein, 2007). Regarding specifically object recognition and affect, information about their temporal order is a potentially crucial piece of evidence for the plausibility of cognitive versus "noncognitive" theories of emotion generation. Cognitive emotion theories assume that emotions presuppose cognitions of their eliciting objects (e.g., Arnold, 1960; Lazarus, 1982; Ortony et al., 1988; Scherer, 2001; see Reisenzein, 2017, for a review). In contrast, noncognitive emotion theories argue that at least some kinds of emotion are at least sometimes, but perhaps even regularly, evoked by stimuli via a "noncognitive" pathway (e.g., LeDoux, 1998; Öhman & Mineka, 2001; Zajonc, 1980).

These contrasting assumptions about the generation of emotions have been most extensively investigated for emotions that are evoked, in a seemingly automatic fashion, by the perception of certain natural objects, such as the pleasant feeling caused by a smiling face, or the unpleasant feeling evoked by a cockroach (e.g., Lang et al., 1993). With regard to these "object-evoked" emotions, a crucial difference between the cognitive and noncognitive theory concerns the question whether or not the evoking objects must be represented *conceptually* to be emotionally effective. Most cognitive emotion theorists answer this question affirmatively (see e.g., Lazarus, 1995; Russell & Woudzia, 1986). As Zajonc (1980, p. 160) put the cognitivist view: "cognitive processes [such as] recognition and categorization are primary" to affect, and hence "before I can like something, I must first know what it is." This hypothesis is called the *semantic primacy* hypothesis of emotion generation (e.g., Lazarus, 1982; Storbeck et al., 2006). More precisely, according to cognitive emotion theorists, the semantic classification of an emotion-evoking object – the (unconscious or conscious) recognition that the object is of a particular type (e.g., a smiling face or a cockroach) – is assumed to be a necessary, partial cause of the emotion. It is only a partial cause, because to evoke an emotion, the object must also be evaluated as positive or negative (e.g., Arnold, 1960; Lazarus, 1995; Ortony et al., 1988). However, here too, the conceptual representation of the object is primary, in that it constitutes the cognitive basis of the evaluation: The object is evaluated as good or bad *because* it is, consciously or unconsciously, recognized as being of a particular kind (e.g., a smiling face or a cockroach).

In contrast, "noncognitive" emotion theorists assume that at least some kinds of emotional experiences do *not* presuppose a conceptual representation of the objects that elicit them, and that one can therefore sometimes like something without (including before) knowing what it is (Zajonc, 1980). This is called the (causal version of the) *affective primacy* hypothesis. In our view, the most plausible explication of this hypothesis is that these emotional experiences are elicited by *pre- or nonconceptual* object representations (see e.g., Crane, 1992; Marr, 1982), such as image-like representations of the prototypical shapes, textures and color patterns of disgusting objects in the case of disgust; or in the case of fear, representations of "sinusoidal shapes related to snakes" (Öhman & Mineka, 2001; see also, LeDoux, 1998).

Empirical Evidence I: Subliminal Perception

The most extensively used experimental paradigm to test the semantic against the affective primacy hypothesis is the subliminal perception paradigm (for a recent summary of this research, see Lähteenmäki et al., 2015). Studies using this paradigm have produced evidence suggesting that subliminally presented (very briefly shown, backward-masked) affective stimuli can bias the evaluation of subsequent stimuli (affective priming; e.g., Murphy & Zajonc, 1993) and can elicit peripheral physiological reactions (e.g., Öhman & Soares, 1994), facial expressions (e.g., Dimberg et al., 2000) and amygdala activity (e.g., Morris et al., 1998) similar to the reactions that are observed if the stimuli are presented above the threshold of conscious perception. Noncognitive emotion theorists haven taken this evidence as support for their hypothesis of a noncognitive path to emotions. This inference is based on the explicit or implicit auxiliary assumption that, since the objects were not consciously perceived, they could not have been semantically categorized, implying also that no categorization-based evaluation could have taken place. The standard response by cognitive emotion theorists has been to accept the empirical findings, but to emphasize that the recognition of objects, as well as their evaluation as good or bad, can occur at an unconscious level of information processing, and that the findings therefore do not refute the cognitive theory of emotion (e.g., Lazarus, 1995; Storbeck et al., 1996).

In addition, however, more recent research suggests that the presentation times used in the subliminal perception studies may not have been short enough to prevent at least partial awareness of the stimuli in a subset of participants or trials (e.g., Lähteenmäki et al., 2015; Pessoa et al., 2005). Recent experiments using a more sensitive psychometric threshold detection procedure, in which the duration of the masked affective pictures was systematically varied from clearly below to clearly above threshold, found that the pictures elicited physiological reactions (e.g., Peira et al., 2012; see also, Codispoti et al., 2009), and could be reliably judged as pleasant or unpleasant (Lähteenmäki et al., 2015) or as evoking a pleasant or unpleasant feeling (Peira et al., 2012), only at presentation times at which the depicted objects could be at least partly recognized.

These findings suggest that the elicitation of emotions may not just require the *recognition* (semantic categorization) of the eliciting objects, but even their *conscious* recognition (for additional supportive evidence, obtained with a variant of the continuous flash suppression paradigm, see Hedger et al., 2015).

Empirical Evidence II: Reaction Time Studies

This conclusion receives further support from the findings of a second, if much less abundant set of studies, in which the latency of object recognition and affect was estimated using speeded reaction time (sRT) tasks.[2] This research goes back to a doctoral student of Edward B. Titchener, Taizo Nakashima (1909a; 1909b), who conducted a series of studies to test an early version of the affective primacy hypothesis (Zajonc, 1980) proposed by Wundt (1896). In three of his experiments, Nakashima used an sRT task to measure how long it takes to report (by a button-press, or vocally) (a) the occurrence of pleasant and unpleasant feelings evoked by simple stimuli (colors, tones, touch stimuli), and (b) the identification of the stimuli (the recognition of the color; the classification of the tones as low, medium, or high; the classification of the touch stimuli as hard or smooth, sharp or blunt etc.). Based on the assumption that causes must precede their effects, the semantic primacy hypothesis predicts that the latency of object recognition is shorter than the latency of affect, whereas the affective primacy hypothesis suggests the opposite temporal relation. In conflict with affective primacy, but supporting semantic primacy, Nakashima (1909a; 1909b) found that these "cognitive" judgments took consistently less time than the judgment of affect onset.

Although Nakashima's studies had a number of limitations when judged by today's methodological standards, their results were supported in a careful set of experiments by Nummenmaa et al. (2010), who found that the median RT for the semantic categorization of affective pictures (e.g., as containing an animal or not) was reliably faster than the categorization of the pictures as pleasant or unpleasant. These findings support Nakashima's conclusions if one assumes, as is likely, that the valence judgments were based, at least for part of the pictures, on the pleasant and unpleasant feelings evoked by the stimuli (cf. Wells, 1925).

Objectives of the Present Research

The studies reported in this article supplement the sRT studies summarized above with data on the subjective timing of object recognition and affect estimated from TOJ and SJ tasks. As noted by Neumann and Niepel (2004), sRT tasks and temporal judgment tasks comprise the two traditional methods of estimating the latencies of subjective experiences, with each method having advantages and disadvantages (see also, Jaśkowski, 2014; Sternberg & Knoll, 1973). Advantages of RT measurements are that they are easy to collect and yield absolute response latencies; their main disadvantage, already emphasized by Nakashima (1909a), is that they also include the time needed for deciding on and executing the response (for discussions, see Fabre-Thorpe, 2011; Jaśkowski, 2014; Miller & Ullrich, 2013; Neumann & Niepel, 2004). For this reason, reaction times are only indirect measurements of perceptual latencies. Furthermore, differences between RTs may overestimate the differences between perceptual latencies (Cardoso-Leite & Gorea, 2010; Miller & Schwartz, 2006).

In contrast, temporal judgments allow the direct measurement of the latency of perceptions and other subjective experiences (Schneider & Bavelier, 2003). Although a disadvantage of temporal judgments is that they only provide estimates of *relative* latencies (i.e., the time difference between a target mental event and the perception

[2]In addition, this conclusion appears to agree with latency estimates of object recognition and affect inferred from event-related potentials (ERPs). However, the interpretation of these data is complicated by the fact that no one-to-one relations between particular ERP components and the mental states of object recognition and affect seem to exist. For a recent review of ERP research on emotions see Hajcak and Foti (2020).

of the probe; Schneider & Bavelier, 2003), for many questions of emotion psychology for which timing information is relevant, including the question of cognitive versus affective primacy, relative latencies are sufficient; in fact, in this and other cases, all one needs to know is the *temporal order* of mental events.

To our knowledge, TOJ or SJ tasks have not been used prior to the studies reported in this article to compare the time needed for object recognition and affect. However, in the course of his pioneering research on the latency of affect, Nakashima (1909a, Exp. 1–3) invented a timing method somewhat similar to the TOJ method to obtain a direct estimate of the latency of pleasure and displeasure. This method – which Nakashima (1909a) for this reason called the "direct reaction method" – was as follows: Participants were shown affective stimuli (colors, geometric figures, pictures of natural objects) for different brief durations, and were asked to report whether they experienced the feeling evoked by a stimulus while it was still visible, or only later. This method can be reconstructed as being a variant of the TOJ procedure in which the probe stimulus consists of the offset of the target stimulus (as a result, stimulus duration and probe SOA are confounded; furthermore, no probe stimuli preceding the target stimulus can be presented). For affective pictures – the stimuli most similar to those used in our experiments – Nakashima's (1909a, Exp. 3) participants (10 trained observers) needed median stimulus presentation times of 390–490 ms (M = 422 ms, SD = 37 ms) to experience a pleasant or unpleasant feeling within the time window of stimulus presentation. However, the perceptual latency of cognitive (recognition) judgments of the stimuli was not measured by Nakashima (1909a) with the "direct reaction method". In addition to using better temporal judgment paradigms and associated analysis methods, this neglect was rectified in our studies.

Hypotheses and Overview

Two hypotheses derived from the semantic primacy hypothesis were tested in our experiments. If, as the semantic primacy hypothesis assumes, object recognition is a partial cause of affect, then (Hypothesis 1) object recognition should temporally precede affect and (Hypothesis 2) not only the *occurrence* of these events, but also *their occurrence times* should be positively correlated (because the later the cause occurs, the later is the effect). A quantitative cognitive model of emotion generation that would allow precise predictions of the size of the expected cognition-affect lag and the correlation does not exist; however, a rough empirical estimate of the expected effect size is possible (see Method of Experiment 1a).

The latencies of object recognition and affect were estimated in three TOJ and one SJ experiment. In Experiment 1a, the latency measurement paradigm is introduced and the latencies of object recognition and affect are estimated with TOJs. Experiment 1b seeks to replicate the findings of Experiment 1a with a broader range of SOAs to rule out the possibility that late affect onsets were overlooked in Experiment 1a. The aim of Experiment 2 is to replicate the findings of Experiments 1a and 1b with the SJ method, regarded by many temporal judgment researchers as less subject to judgment biases than the TOJ method. Experiment 3, another TOJ experiment, aims to generalize the findings to a much larger and more heterogeneous set of affective pictures, as well to an experimental condition where each picture is presented only once and thus less familiar, and where pleasant and unpleasant pictures are intermixed. The last procedural change addresses a potential objection to the blockwise presentation of the pleasant and unpleasant pictures in the preceding experiments (see Experiment 3). Finally, the results of the four experiments are meta-analytically integrated to obtain cross-experiment and cross-method estimates of the size of the cognition-affect lag and the correlation between the latencies of cognition and affect, and to profit from the higher precision of the combined estimates and the increased power of the associated statistical tests (see Schimmack, 2012). In addition, the possible effect of affect intensity on the latency of affect is analyzed. The analyses focus on the PSS, but we also exploratively analyze another common performance measure of psychophysical tasks, the JND (just noticeably difference).

Ethics Statement

All experiments were designed and conducted in accordance with the Code of Ethics of the World Medical Association ("World Medical Association Declaration of Helsinki," 2013). All participants in the four experiments gave their informed consent to participate prior to the start of the experiments. They were informed that they would see pleasant and unpleasant pictures and that they could terminate their participation at any time without negative consequences. Ethical review and approval were not required for these studies in accordance with the German legislation and with institutional requirements. In the following description of our experiments, we report all measures, manipulations, and exclusions (if any). The studies were not preregistered. All analyzed data and R scripts are available on https://osf.io/eugc5/.[3]

Experiment 1a: Temporal Order Judgments

Method

Participants

The participants were 20 introductory psychology students, 15 females and five males (M age = 21.2, SD = 3.4), who participated as part of their study requirements. Two participants were excluded from the data analyses because they did not meet our inclusion criteria described below, reducing the final sample size to N = 18.

The planning of the sample size focused on the central Hypothesis 1. Due to the absence of previous comparable studies using temporal order judgments, a precise estimate of the effect size of the expected time difference between object recognition and affect was not available. However, Nakashima's (1909a; 1909b) three sRT experiments on the latency of cognitive and affective judgments to colors, tones and touch stimuli, which come closest to our studies in terms of the affective judgment made by the participants ("When did you notice the feeling evoked by the picture?"), suggest a large effect (generalized η^2 = .44; Bakeman, 2005; calculated by us from

[3]Because Experiment 3 was conducted during the Sars-CoV-2 epidemic in 2021, strict measures were taken to minimize the risk of infection in the lab version of this experiment: Only one subject participated at a time; the participants underwent antibody testing for the corona virus in a separate room prior to the experiment; the experimenter was not present in the room during the experiment; participants and experimenter wore FFP2 face masks; their interaction was limited to a brief introduction at the beginning and participant remuneration at the end of the experiment; the room was aired out between sessions for half an hour; and the participant's work place was disinfected after each session.

the published median RTs of the 12 participants). Nummenmaa et al. (2010) also obtained large effect sizes for the RT latencies of semantic versus affective judgments (e.g., $\eta_g^2 = .30$ in Exp. 2 and .29 in Exp. 3).[4] We therefore expected to obtain effects of a similar size. This expectation was supported by the results of our experiments, which found a combined effect size (the sample-sized weighted mean of the effect size found in each study) of $\eta_g^2 = .19$ To detect an effect of this size in the within-subject design used in our studies at $\alpha = .05$ with a power of .80, assuming a correlation of .51 between the two onset times (corresponding to that actually obtained; see the meta-analytic summary of results), a sample size of 11 is sufficient (calculated with G*Power; Faul et al., 2007).

Design

The experiment had a 2 (task: object recognition or affect) × 2 (picture valence: pleasant or unpleasant) × 2 (task order: object-recognition TOJ or affect-experience TOJ first) design. The first factor was varied within subjects and the remaining two factors between subjects. Task order was included to control for possible sequence effects. Picture valence was varied between subjects to avoid overtaxing the participants' attention and motivation (a within-subject manipulation would have doubled the length of the experiment) and to make the affect task as simple as possible: Because affect of a single quality was induced, the participants could focus on the detection of affect of a known quality (pleasure or displeasure/disgust). Participants were randomly assigned to the four groups resulting from the crossing of the two between-subject factors.

Procedure

The experiment was programmed using DMDX (Forster & Forster, 2003). It comprised four consecutive tasks: an affect intensity rating, three blocks with different kinds of TOJ, and a repetition of the affect rating. All tasks were completed in a single session in a computer lab in groups comprising 2–4 subjects. Stimuli were presented on TFT monitors with a resolution of 1280 × 1024 pixels and a refresh rate of 75 Hz. Work places were separated by room dividers and each participant worked at his or her own pace. The experiment took about half an hour.

Affective Stimuli. Twelve pleasant and 12 unpleasant pictures from a previous study (Junge & Reisenzein, 2016, Exp. 1) were used. The pleasant pictures showed among others a laughing child, a bunny, a kitten, a panda bear, a sunflower field and an ice cream cup; the disgusting pictures showed among others a purulent finger, vomit, a snake pit, maggots, a moldy piece of bread and an overflowing ashtray. We focused on disgust as the unpleasant feeling because disgust is an emotion for which the hypothesis of nonconceptual affect elicitation seems particularly plausible. The reason is that disgust-elicitors comprise a limited number of categories of objects with characteristic appearances (e.g., Curtis et al., 2004). In addition, the evolutionary argument for an adaptive advantage of fast affective reactions, typically illustrated with fear (LeDoux, 1998; Öhman & Mineka, 2001) also seems plausible for disgusting

[4] We computed these effect sizes as follows: We first estimated d from the published within-subject t-tests, using formula 3 in Dunlop et al. (1996), assuming a correlation of .43 between the RTs of semantic and affective judgments, as obtained by Nummenmaa et al. (2010) in their Experiment 3 (although for the by-pictures analysis). Subsequently, d was transformed into η^2.

stimuli: Not getting into contact with contaminating or infectious substances should also benefit from fast responding.

The pictures were 300 pixels wide and 360 pixels high and were presented at the center of the monitor at normal viewing distance. Junge and Reisenzein (2016) found that the pictures elicited pleasant feelings and disgust, respectively, of medium to high intensity in most participants. Previous research has found that comparable pleasant and unpleasant pictures also tend to evoke weak physiological reactions and facial expressions (activation of the zygomaticus vs. corrugator muscles) in viewers (e.g., Codispoti et al., 2006; Lang et al., 1993).

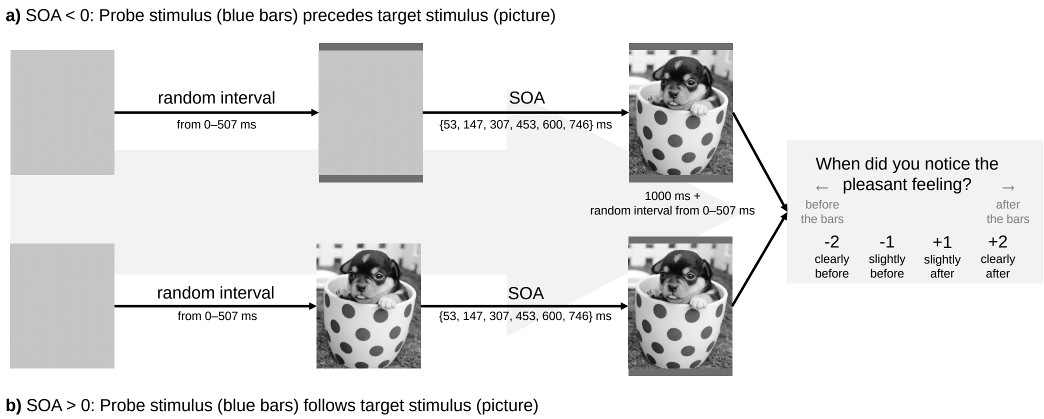
Affect Rating Task. Prior to the TOJ tasks, the (depending on group) pleasant or disgusting pictures were separately presented to the participants, who were asked to rate how pleasant or unpleasant they experienced each picture to be on a scale ranging from 0 = "not at all pleasant [unpleasant]" to 10 = "extremely pleasant [unpleasant]". The affect ratings served to familiarize the participants with the pictures and to allow us to verify that they elicited at least moderately intense feelings in each participant. To test whether the pictures retained their affect-evoking potential until the end of the experiment, the ratings were repeated after the TOJ tasks.

TOJ Tasks. To familiarize the participants with the TOJ procedure, they first completed a training block in which pictures showing one of six plain colors (red, green, blue, yellow, black, white) were presented and the participants indicated when they recognized the color: Before or after they noticed the probe stimulus. Subsequently, they completed the object recognition (OR) and affect (AF) TOJ, with task order counterbalanced between subjects. Half of the subjects judged the pleasant and the other half the unpleasant pictures. Each picture was shown once together with the probe stimulus temporally displaced by the 12 SOAs (see below); hence each participant made 12 (pictures) $\times$ 12 (SOAs) = 144 OR TOJ trials and 144 AF TOJ trials. The recognition judgments thus referred to objects that the participants had already seen before. This was meant to simulate typical everyday situations in which familiar affect-evoking objects are encountered (e.g., seeing one's cat repeatedly). The six pictures used in the color-recognition TOJ were shown twice at each probe SOA, resulting in 24 training trials.

TOJ Trial Structure. In each trial, a picture was presented that was preceded or followed, at one of 12 SOAs, by two blue bars (300 $\times$ 15 pixels) shown above and below the image. The perception of the bars served as the probe event whose perceived onset was compared, in each block, with that of different target mental events evoked by the pictures. Specifically, the participants were asked to indicate when they recognized the color of the picture (training block), recognized the object shown in the picture (object-recognition TOJ), and noticed the affect evoked by the picture (affect TOJ): Before, or after the appearance of the probe stimulus. We assumed that the detection of the probe would need comparatively few attentional resources, allowing the participants to focus on OR and AF. The wording of the instruction, which drew attention to the target events, was chosen to support this attentional focus.

In each block of trials, the target-probe (picture-bar) combinations were presented in an individually randomized order. To avoid the development of precise expectations about the onsets and durations of the stimuli, the first stimulus appeared after a pseudo-random interval of 507 to 1000 ms following trial onset, which was signaled by the appearance of a uniformly gray picture. After one of 12 SOAs, the second stimulus was added. The two stimuli then remained on the screen for a fixed presentation time of 1000 ms plus another random interval of 507 to 1000 ms (see Figure 1).

Figure 1
Stimulus Presentation in the TOJ Task (Experiment 1a)



Note: a) Shows trials with negative SOAs, meaning that the probe stimulus (blue bars; here shown in dark gray) preceded the target stimulus (picture); b) shows trials with positive SOAs, meaning that the probe stimulus followed the target stimulus. The probe in a) and the target in b) appeared after a pseudo-random interval between 507 and 1000 ms following the onset of a uniformly gray picture. After the SOA interval, the picture-plus-bar combination remained on the screen for 1000 ms plus another random interval of 507 to 1000 ms. After this final interval, the response screen was presented. To avoid possible copyright issues, we have replaced the original example image by a picture with comparable affective quality from the *Open Affective Standardized Image Set* (OASIS; Kurdi et al., 2016) used in Experiment 3.

Following each picture-bar presentation, the screen was cleared and a temporal order question appropriate to the TOJ task was presented at the center of the monitor (see Figure 1). That is, the participants were asked, when – before or after the occurrence of the blue bars – they had identified the color of the picture, had recognized the object or objects shown in the picture, or had noticed the (depending on condition) pleasant or unpleasant feeling evoked by the picture. These questions were identical (AF) or conceptually identical (OR) to those used by Nakashima (1909a; 1909b; see also Titchener, 1905, for the OR question). In particular, we followed Nakashima (1909a) in using the AF question "When did you notice the pleasant [unpleasant] feeling?" rather than the possible alternative question "When did you notice that the picture was pleasant [unpleasant]?" (e.g., Nummenmaa et al., 2010; Wells, 1925) to make sure that the participants focused on the feelings evoked by the stimuli, rather than on object evaluations, which could have been retrieved from memory (Fazio, 1995; Wells, 1925). This issue is discussed in more depth in Franikowski and Reisenzein (in press).

Following each TOJ question, a response scale with four categories, labeled "-2 clearly before", "-1 slightly before", "+1 slightly after", and "+2 clearly after", was presented. A graded response scale was used to encourage the participants to make fine-grained temporal judgments (cf. Junge & Reisenzein, 2013) in an attempt to reduce possible judgment biases that may occur in the TOJ task when participants are uncertain about the temporal order (cf. García et al., 2013; Linares & Holcombe, 2014; Schneider & Bavelier, 2003). The two "before" and "after" judgments were subsequently collapsed into a "before" and "after" category, respectively, to be able to fit the psychometric model typically used in TOJ research. The participants entered their answers using the two keys immediately to the left and right of the spacebar (labeled -2, -1 and +1, +2 respectively). To reduce possible reporting errors (cf. García-Pérez & Alcalá-Quintana, 2012; Wichmann & Hill, 2001), the participants were instructed to use the index and middle finger of their left and right hand, respectively, for the left and right response keys, and to keep the fingers close to the keys during the experiment.

Selection of SOAs. The probe stimulus was presented at six SOAs preceding the affective pictures (-746, -600, -453, -307, -147, -53 ms) and at six SOAs following the pictures (53, 147, 307, 453, 600, 746 ms). To allow exact picture onsets, all SOAs were multiples of the refresh rate (13.33 ms) of the 75 Hz monitor. The three intervals closest to the center were kept smaller than the remaining intervals to permit better temporal resolution at these intervals (Kingdom & Prins, 2010). The maximal SOA of 746 ms was extrapolated from the findings of previous research, which suggests that pleasant and unpleasant feelings evoked by pictures can be detected within 500 ms (Nakaskima, 1909a, Exp.3) and can be reported by a key press within 800 ms (Wells, 1925) and by an eye-gaze, within 350 ms (Nummenmaa et al, 2010, Exp. 3; assuming that the valence judgments made in these experiments were based on experienced affect). Given these estimates, the time window in our experiment seemed broad enough to capture even comparatively late onsets of affect.

However, because the psychometric functions of one participant of Experiment 1a and one from Experiment 2 raised the possibility that their PSS for affect exceeded 746 ms, we also conducted, following these experiments, a control TOJ experiment with 13 participants from the same subject pool and with the same stimuli, but a much broader time window (-1800 to +1800 ms). This experiment, reported below as Experiment 1b, confirmed that the SOA range used in Experiments 1a and 2 was sufficient.

Data Analysis

The data analysis of the TOJ tasks followed a psychophysical analysis protocol recommended e.g., by Spence et al. (2003), that comprises three steps: The fitting of psychometric functions to the data of the individual participants; the extraction of performance measures (PSS and JND) from the fitted curves, and the analysis of the performance parameters using ANOVA (analysis of variance).

Fitting Psychometric Functions. To analyze the TOJ data, the "target event before probe" judgments were recoded as 1 and the "target after probe" judgments as 0. Following the classical approach to the estimation of S-shaped psychometric functions (e.g., Knoblauch & Maloney, 2012; Wichmann & Hill, 2001), that is also typically used in TOJ research (e.g., Linares & Holcombe, 2014; van Eijk et al., 2008), we then fitted a cumulative normal distribution function (i.e., a probit model; cf. Hardin & Hilbe, 2007; Knoblauch & Maloney, 2012) to the proportions of "target event before probe" answers obtained at the different SOAs. We fitted an extended version of the probit model that includes offsets to the lower and upper asymptote of the psychometric function to account for possible lapses of attention, response errors (pressing the wrong key) and other kinds of errors (see Linares & Holcombe, 2014; van Eijk et al., 2008). That is, the estimated psychometric function was

$$p(x) = \gamma + (1 - \gamma - \lambda)\Phi(x; \mu, \sigma)$$

where p(x) is the probability of the "target event before probe" response, x is the delay time, Φ is the cumulative distribution function of the normal distribution with mean μ and standard deviation σ, γ is the lower asymptote and λ the distance between 1 and the upper asymptote (e.g., Knoblauch & Maloney, 2012; Wichmann & Hill, 2001).

Although this psychometric function, or a restricted version of it (without γ and λ), is often fitted to TOJ data without making explicit assumptions about the underlying judgment process (cf. García-Pérez & Alcalá-Quintana, 2012), it is consistent with the following simple model of that process (e.g., Schneider & Bavelier, 2003; Yarrow et al., 2011): (1) the response "the target event occurred before the probe" is given if a latent decision variable, which represents the perceived difference between the occurrence times of the two mental events, exceeds a fixed decision criterion; (2) the time differences for each SOA are normally distributed with constant variance; and (3) even at high negative and positive SOAs (i.e., even if the probe stimulus is presented much earlier or later than the target stimulus), participants can make errors with probabilities γ and λ, respectively, because they miss the onset of the target or probe event, confuse the response buttons, etc. As a consequence, the psychometric function asymptotes at γ and $(1 - \lambda)$ rather than 0 and 1.

The four-parameter psychometric function was estimated using maximum likelihood. Estimation was carried out in R (R Core Development Team, 2018) using the add-on package *quickpsy* (Linares & López-Moliner, 2016). The estimations were performed separately for each participant and task (OR < AF). To guard against local minima, an attempt was made in each case to further improve the obtained fits by restricting the initially used, broad parameter bounds (in particular σ; cf. Linares & López-Moliner, 2016); however, in most cases, the initial solution was already optimal. In these analyses, we relied on the BIC (Bayesian Information Criterion) as a relative measure of fit. However, to describe the final fit, we use – again following some previous TOJ research (e.g., Love et al., 2013) – R^2 as a well-understood measure of absolute fit (on the merits of R^2, see Nakagawa & Schielzeth, 2013; Zheng & Agresti, 2010).

PSS and JND. The estimated psychometric functions were used to determine, for each participant and task (OR < AF), the point of subjective simultaneity (PSS), defined as the estimated time point at which "before" and "after" judgments would be equally likely (.50). While our main interest was on the PSS, we also computed a second common performance measure, the JND (just noticeable difference), defined as the difference on the x-axis (SOA) between the .75 and .50 value of the psychometric function. The JND is a direct function of the spread (σ) of the psychometric curve and is standardly interpreted as a measure of the observer's sensitivity – the smallest difference on the stimulus scale (in the TOJ task, the smallest interval between target and probe) that the observer can reliably detect. However, different from the classical temporal order experiments that use simple stimuli with constant perceptual latencies (up to random error), we fitted the psychometric function to the participant's responses to 12 different pictures, each with a somewhat different latency of OR and AF (evidence for this was obtained in the by-pictures analysis reported below). Therefore, the estimated PSS of OR and AF of each participant is actually an average of the participant's perceptual latencies for the different pictures, and the spread of the psychometric curves indexed by the JND reflects not only the observer's sensitivity, but also differences in the actual occurrence times of OR and AF for the 12 pictures (see e.g., Estes, 1956).

Both performance measures (PSS and JND) were computed from the estimated psychometric function after the error parameters (γ and λ) had been removed (i.e., had been set to 0). As a consequence, the PSS and JND refer to the latent psychometric function Φ in the above equation, which (if the model is correct) describes the participant's perception of the temporal relations between target event and probe for the different SOAs uncontaminated by performance errors (cf. van Eijk et al., 2008; García-Pérez & Alcalá-Quintana, 2012). In contrast to the (typical) shape of the complete psychometric function, Φ ranges between 0 and 1 on the y-axis, is symmetric, and the PSS for Φ corresponds to μ.

The individual performance measures were processed further using mixed-design ANOVAs. To protect against the accumulation of α error in multiway ANOVA, the p-values of all effects except those for the predicted effect of task (OR or AF) on the PSS were adjusted using the Holm procedure (Cramer et al., 2016). In addition to the by-subjects analysis, we also performed, following Nummenmaa et al. (2010), a by-pictures analysis. All correlations were tested two-tailed.

GPAQ Model. In recent years, more sophisticated psychometric models for TOJs and SJs have been proposed that allow fitting asymmetric psychometric curves and promise to provide more accurate estimates of the perceptual latencies (see in particular, García-Pérez & Alcalá-Quintana, 2012; Tünnermann et al., 2017). However, these models require high-precision data to obtain reliable estimates of all parameters, which in practice means to collect a large number of trials per SOA. For example, Alcalá-Quintana and García-Pérez (2013) suggest > 20 trials per SOA for TOJ tasks, and about 20 for SJ tasks. With 12 trials per SOA, our data did not meet this requirement. In addition, the judgments required of our participants were more complex and therefore noisier than those required in typical TOJ and SJ studies. However, we used the psychometric model proposed by García-Pérez and Alcalá-Quintana (2012; henceforth the GPAQ model) to check the robustness of our PSS and JND estimates, taking advantage of the potential of the GPAQ model to fit the data more closely than the probit model. The same was done for the SJ data collected in Experiment 2.

Data exclusions. We decided a priori to exclude participants who gave zero or very low affect ratings to most pictures, because it is not meaningful to measure the onset of feelings if no feelings are evoked. In addition, we decided to follow common practice in the psychometric modeling of TOJ and SJ data (see e.g., van Eijk et al., 2008; Kostaki & Vataki, 2018; Love et al., 2011; Spence et al., 2003) to exclude participants whose data did not allow estimation of the PSS (without extrapolating beyond that SOA range), or suggested random or task-irrelevant responding (e.g., low curve fits; the same response given at nearly all SOAs). As reported in Supplemental Materials S1, highly similar results were obtained if all participants were included.

Results

Inspection of the data of Experiment 1a revealed that two participants did not meet our inclusion criteria. One participant, in the negative affect group, gave zero or very low "unpleasant" ratings to most pictures, suggesting that the pictures failed to elicit disgust. The second participant, from the pleasant affect group, was excluded because her psychometric function for affect did not approach the lower and upper asymptote even at the longest post-picture SOAs (±746 ms), resulting in a bad curve fit. Hence, the final sample for the TOJ task comprised 18 participants. As reported in Supplementary Materials S1, the average PSSs and JNDs changed only minimally if these participants were included. For one of the two excluded cases, the PSS for OR and AF was in line with predictions (OR < AF).

Affect Ratings

The mean pleasantness ratings of the pleasant pictures were 7.8 before and 7.4 after the TOJ tasks; the mean unpleasantness ratings of the unpleasant pictures were 6.3 and 6.1. Repeated-measures ANOVAs revealed, for both kinds of pictures, a significant effect of picture identity, Fs $\geq$ 3.30, ps $\leq$ .002, $\eta^2_g \geq .17$ (meaning that the pictures differed in the intensity of affect that they evoked), whereas time of measurement and the interaction were not significant, ps $\geq$ .257. In addition, the participants' ratings of the pictures at the two measurement points were highly correlated (mean intraindividual after a Fisher r-to-z transformation r = .84 for the pleasant and r = .83 for the unpleasant pictures). These findings imply that the pre- and post-task ratings were very similar. We therefore used their average as an overall index of the intensity of affect evoked by the pictures. On this index, the average intensity of pleasure caused by the 12 pleasant pictures ranged from 5.5 to 8.9 on the 0 to 10 scale (M = 7.6, SD = 0.9), and that of displeasure evoked by the unpleasant pictures, from 2.7 to 8.6 (M = 6.2, SD = 1.8).

For individual participants, the average rating of the pleasant pictures ranged from 6.2 to 9.7 (M = 7.6, SD = 1.0) and that of the unpleasant pictures from 5.0 to 7.5 (M = 6.2, SD = 0.9). All participants rated at least 9 of the 12 pleasant pictures and 7 of the 12 unpleasant pictures > 4 on the 0–10 scale.

Taken together, these findings suggest that the pictures evoked pleasant and unpleasant affect of at least moderate intensity in all included participants, and that the repeated presentation of the pictures during the TOJ tasks caused only minimal affective adaptation on the subjective level. Both conclusions agree with previous findings (Junge & Reisenzein, 2016; see also, Codispoti et al., 2006).

Temporal Order Judgments

Table 1 summarizes the individual fits (R^2) of the probit model to the OR and AF TOJs, as well as the PSSs and JNDs computed on the basis of the latent psychometric functions (i.e., after removing the error parameters γ and λ).

Model Fits. The average fit of the probit model to the TOJs was high (mean R^2 = .97, SD = .03 for OR, and .95, SD = .03 for AF). Figure 2 shows the psychometric function for the mean parameters of the included participants after removal of the error parameters; this curve represents the mean latent temporal discrimination

Table 1
Fit Index (R^2), PSS and JND for the TOJ Task in Experiment 1a

	R^2			PSS			JND		
Task	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Object recognition	.97	.03	[.88, >.99]	42	69	[-106, 119]	59	44	[2, 186]
Affect	.95	.03	[.88, .99]	121	86	[-18, 246]	80	43	[8, 144]

Note: $N = 18$.

performance at different SOAs. The psychometric functions of all 20 individual participants are contained in Supplementary Materials S1 (Figure S1.1).

PSSs of OR and AF. Hypothesis 1 holds that the latency of OR is faster than that of AF. To test this hypothesis, a mixed-design ANOVA was conducted on the latencies, with type of judgment (OR vs. AF) as the within-subject factor and the control factors picture valence (pleasant vs. unpleasant) and task order (OR TOJ vs. AF TOJ second) as between-subject factors. In support of Hypothesis 1, this analysis revealed a significant effect of task, $F(1, 14) = 12.80$, $p = .003$, $\eta^2_g = .25$. No other effect reached significance, $F\text{s} \leq 3.29$, $p\text{s} \geq .546$. As can be seen from Table 1 and Figure 2, the mean PSS was 42 ms for OR and 121 ms for AF, meaning that compared to the time needed to detect the onset of the probe (blue bars), the participants needed on average 42 ms more to recognize the object shown in the pictures, and 121 ms more to notice the affect evoked by the pictures. Hence, they became on average aware of the feelings 79 ms after they recognized the object. Examination of the individual PSSs revealed that 15 of the 18 participants agreed with the mean pattern (PSS AF > PSS OR); the remaining four showed the opposite pattern, with PSS differences of -11, -13 and -114 ms, respectively. As explained in the General Discussion, we believe that these participants found the affect task too difficult and switched to a simplifying response strategy.

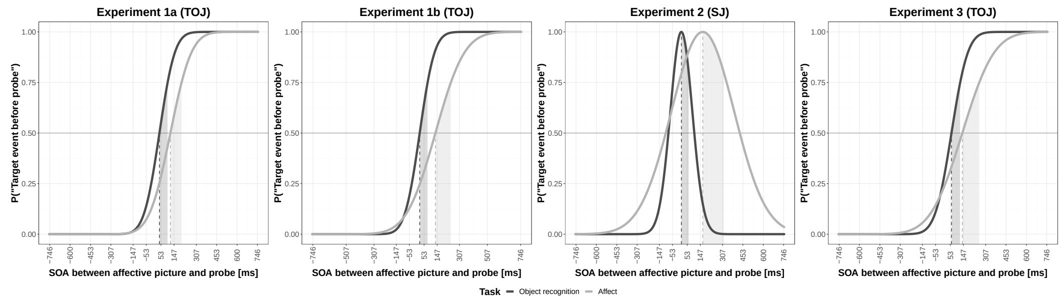
Correlation Between the Latencies of OR and AF. Hypothesis 2 holds that the latencies of OR and AF are positively correlated. Although the correlation was positive as predicted, $r(16) = .26$, it was not significant given the sample size, $p = .294$ (two-sided test). Hence, Hypothesis 2 received at best limited support.

JNDs. As shown in Table 1, the spread of the psychometric function (JND) was descriptively greater for AF ($M = 80$ ms) than OR ($M = 59$ ms); however, an ANOVA analogous to that for the PSS revealed that this difference did not reach significance after the Holm correction, $F(1, 14) = 2.54$, $p = .933$. The other effects were also not significant, $F\text{s} \leq 1.24$, $p\text{s} \geq .826$.

Replication With the GPAQ-TOJ Model. To check to which degree the PSS and JND estimates depended on the estimation procedure, we also fitted the psychometric model for TOJs proposed by García-Pérez and Alcalá-Quintana (2012; Alcalá-Quintana & García-Pérez, 2013) to the data. As reported in Supplementary Materials S1, the GPAQ model fitted the data significantly better than the probit model. The main reason was that the psychometric curves were often somewhat asymmetric, which was captured by the GPAQ but not the probit model. Nevertheless, the average PSSs (OR: 58 ms, AF: 117 ms) and JNDs (OR: 65 ms, AF: 117 ms) were very similar to those obtained with the probit model (Table 1; for a direct

Figure 2

Psychometric Functions for Object Recognition and Affect, Fitted to TOJ (Experiments 1a, 1b, and 3) and SJ (Experiment 2) data



Note: The graphs show the average latent psychometric function, $\Phi(x;\, \mu,\, \sigma)$, obtained by averaging the parameters for location (μ) and spread (σ) estimated for the individual participants; the error parameters (γ and λ) are removed in these functions (i.e., set to 0). The vertical dashed lines indicate the PSSs for object recognition (dark-gray line) and affect (light-gray line); the shaded areas indicate the JNDs for the TOJs and a comparable JND value calculated from the HWHH for the SJs (see text for an explanation). The SOA range (x-axis) used in experiments 1b and 3 has been trimmed to the range used in the experiments 1a and 2 (-746 ms to +746 ms) to make all curves visually comparable. Note that all psychometric curves reached their asymptotes within this range.

comparison, see Table S1.2 in Supplementary Materials S1) and the correlation between the individual PSS estimates obtained with the probit and GPAQ model was r = .85 for both OR and AF.

Analysis for Pictures as Units. Following Nummenmaa et al. (2010), we also conducted a by-pictures analysis; that is, we fitted the probit model to the TOJs given to each of the 24 individual pictures (12 pleasant, 12 unpleasant), with data pooled across participants. This analysis provides information about the homogeneity of the responses to the different pictures, although only at the group level. Because the pleasant and unpleasant pictures were judged by different groups, these analyses were only based on 10 (pleasant pictures) or 8 (unpleasant pictures) data points (judgments) per SOA. Despite the low N, high fits were obtained for all pictures (mean R^2 = .97, SD = .02 for OR and .93, SD = .04 for AF), and the average PSS obtained for OR (47 ms, SD = 32 ms) and AF (123 ms, SD = 53 ms) was nearly identical to those obtained for participants as the units of analysis, implying that the cognition-affect delay was also nearly the same (76 ms). An ANOVA for task and picture valence revealed a significant effect of task, $F(1,\, 22) = 43.99$, $p <$ .001, $\eta^2_g = .44$, whereas picture valence and the interaction of both effects were not significant, Fs ≤ 0.22, ps $\geq .769$. For 22 of the 24 pictures, the PSS of AF occurred later than the PSS of OR. Paralleling the by-subjects analysis, the ANOVA for the JND revealed no significant effects, Fs ≤ 5.13, ps $\geq .101$.

Discussion

The central finding of Experiment 1a is that the participants became aware of the affect evoked by pleasant and unpleasant pictures significantly later than they recognized the object shown in the pictures, with an average delay of 79 ms. The effect of task was independent of the order in which the OR and AF judgments were made. This means, in particular, that the shorter perceptual latency of OR cannot be attributed to a training effect (as would be conceivable in the affect-first group, because the participants had already seen the pictures repeatedly when they performed the OR task). In addition, the effect of task did not significantly depend on the hedonic quality of the affect (pleasure vs. displeasure), and the effect was also found when pictures were used as the units of analysis, with nearly the same estimated delay (76 ms). Furthermore, a positive lag between OR and AF was found for most participants (15 of 18) and pictures (22 of 24).

Experiment 1b: Temporal Order Judgments With a Broader SOA Range

As mentioned in the *Method* section of Experiment 1a, we conducted a control study to make sure that the SOA range used in Experiment 1a and Experiment 2 was broad enough to capture even late affect onsets. As this study is a "close" to "very close" replication (LeBel et al., 2018) of Experiment 1a, it is reported here.

Method

Fourteen participants from the same subject pool as in Experiment 1 participated (six females and seven males, M age = 25.4, SD = 4.1). Data from one participant were lost due to computer malfunction, reducing the sample size to 13. All 13 participants met our inclusion criteria. The stimuli and procedure were the same as in Experiment 1a with the exception that an SOA range from -1800 to 1800 ms and eight rather than six positive and negative SOAs were used, and that only one task order (OR–AF) was used because no evidence for an effect of task order was found in Experiment 1a; but note that even if the perceptual latencies could be accelerated by training, the chosen task order works against the hypothesis of a cognition-affect delay. The three smallest SOAs were identical to the smallest SOAs used in Experiment 1a (± 53, 147, 307 ms); the remaining SOAs were ± 507, 746, 1053, 1400, and 1800 ms. The SOAs were unevenly spaced to obtain higher sensitivity for small SOAs (cf. Kingdom & Prins, 2010).

Results and Discussion

Affect Ratings

The analysis of the affect ratings yielded findings similar to those obtained in Experiment 1a. Most important, the pictures induced at least moderately intense pleasant or unpleasant affect in all participants, and largely retained their affect-inducing capacity from the pre-TOJ rating (positive: M = 7.7, negative: M = 5.5) to the post-TOJ rating (positive: M = 7.5, negative: M = 5.2).

Temporal Order Judgments

Model Fits. Model fits were similar to those obtained in Experiment 1a (see Table 2). The estimated psychometric functions for the individual participants can

Table 2
Fit Index (R^2), PSS and JND for the TOJ Task in Experiment 1b

	R^2			PSS			JND		
Task	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Object recognition	.96	.03	[.87, .99]	21	85	[-95, 216]	57	50	[10, 159]
Affect	.94	.06	[.83, > .99]	134	132	[-37, 340]	111	74	[10, 250]

Note: $N = 13$.

be found in Supplementary Materials S1 (Figure S1.2). The latent (error-free) curve for the averaged parameters is shown in Figure 2.

PSSs for OR and AF. Supporting Hypothesis 1, the average PSS of OR found in Experiment 1b was again earlier (21 ms) than the average PSS of AF (134 ms), $F(1, 11) = 8.68$, $p = .013$, $\eta^2_g = .24$ (see also Table 2). The two PSS estimates are in fair agreement with those obtained in Experiment 1a (Table 1), as is the delay between OR and AF (113 ms). Twelve of the 13 participants conformed to the mean pattern (OR < AF). Picture valence and the interaction were not significant, $F\text{s} \leq 1.13$, $p\text{s} \geq .469$. Hence, Hypothesis 1 (OR < AF) was again supported.

Correlation Between the Latencies of OR and AF. The correlation between the latencies of OR and AF was again positive as predicted (Hypothesis 2), $r(11) = .19$, but not significant, $p = .534$ (two-sided test) given the sample size.

JNDs. As in Experiment 1a, the spread of the psychometric function (JND) was greater for AF (111 ms) than for OR (57 ms); however, again, this difference did not reach significance after the Holm correction, $F(1, 11) = 6.55$, $p = .080$. Picture valence and the interaction were also not significant, $F\text{s} \leq 0.50$, $p\text{s} \geq .796$.

Analysis for Pictures as Units. Most likely because of the low number of participants per picture (6 for pleasant and 7 for unpleasant pictures), the model fits were lower than in Experiment 1a, being on average $R^2 = .92$ for OR ($SD = .05$) and $R^2 = .89$ for AF ($SD = .07$). Nonetheless, the by-pictures analysis confirmed the lower PSS for OR (20 ms) compared to AF ($M = 141$ ms), $F(1, 22) = 20.02$, $p < .001$, $\eta^2_g = .33$. The JND was lower for OR ($M = 100$ ms) than AF ($M = 186$ ms), and this effect reached significance in the ANOVA, $F(1, 22) = 8.02$, $p = .029$, $\eta^2_g = .15$. In addition, there was a significant effect of valence, $F(1, 22) = 6.84$, $p = .032$, $\eta^2_g = .14$, but no significant interaction, $F(1, 22) = 2.19$, $p = .153$. Twenty of the 24 pictures corresponded to the PSS pattern obtained for the means (OR < AF). In sum, the findings of Experiment 1a were replicated.

Experiment 2: Simultaneity Judgments

To test whether the findings of Experiments 1a and 1b can be replicated with a different temporal judgment task, Experiment 2 used simultaneity judgments (SJs) to estimate the onset times of OR and AF. The SJ task is regarded by many temporal judgment researchers as superior to the TOJ task in being less subject to judgment biases resulting from the setting of different response criteria (e.g., García-Pérez & Alcalá-Quintana, 2012; Schneider & Bavelier, 2003; van Eijk et al., 2008; Yates &

Nicholls, 2011). However, as Linares and Holcombe (2014) point out, the SJ task is likely also subject to biases, although to different ones than the TOJ task; therefore, trust in the results obtained with each task increases if similar results are obtained with the other task.

Method

Participants

Twenty individuals from the same subject pool as in the previous experiments participated, 15 females and 5 males (M age = 21.4, SD = 2.8). Four participants were excluded from the data analyses because they did not meet our inclusion criteria (see Results); hence the final sample size was N = 16.

Design and Procedure

The design, materials and procedure of Experiment 2 were the same as in Experiment 1a with two differences. First, the participants made simultaneity judgments instead of temporal order judgments. Second, the SOA range for OR was shortened to obtain a more precise PSS estimate.

As in the TOJ experiments, each trial consisted of the presentation of a picture that was preceded or followed, at one of 12 different SOAs, by the probe stimulus (blue bars). However, in the SJ task, the participants indicated whether the bars appeared vs. did not appear at the same time when they (a) identified the color of the picture (color-identification SJ), (b) recognized the object shown in the picture (OR SJ) or (c) noticed the pleasant or unpleasant feeling evoked by the picture (AF SJ).

Experiment 1a and 1b had found that OR took only 20–40 ms more than the perception of the probe, and that the object was recognized with near certainty at an SOA of about 200 ms (see Figure 2). As a consequence, longer SOAs did not provide much information about the PSS of OR. Therefore, in Experiment 2, the SOAs for the OR task were restricted to range from -200 to +200 ms; the exact times were: ± 27, 53, 80, 107, 147, and 200 ms. The same SOAs were also used in the practice task (color identification). For the affect SJ task, the same time intervals as in Experiment 1a were used, ranging from -746 to +746 ms.

Data Analysis

Analogous to Experiments 1a and b, a psychometric function was fitted, separately for each participant, to the proportion of "simultaneous" judgments obtained for the 12 SOAs separating target and probe stimulus. Following a common approach in SJ research (e.g., Fujisaki et al., 2004), we fitted a Gaussian function to the data. This function is usually not motivated by a judgment model but justified on phenomenological grounds, i.e., by the finding that the typical shape of the empirical psychometric curve of SJ data resembles a Gaussian (bell-shaped) function that peaks at the point of subjective simultaneity (see Figure 2). However, judgment models for SJs consistent with this psychometric function have been proposed (see Schneider & Bavelier, 2003; Yarrow et al., 2011; Yarrow, 2018). To stay as close as possible to the TOJ estimation procedure, we fitted the following four-parameter version of the Gaussian function:

$$p(x) = \gamma + (1 - \gamma - \lambda)\Phi(x; \mu, \sigma), \text{ with } \Phi(x; \mu, \sigma) = e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Here, p(x) is the probability of the “simultaneous” response, x is the delay time, μ is the mean and σ the standard deviation of the Gaussian function, γ is the lower asymptote of the curve, and λ is the distance of its peak (the maximum proportion of “simultaneous” judgments) to 1. Analogous to the γ and λ parameters of the extended probit model for TOJs, γ and λ can be interpreted as capturing lapses of attention and response errors for “non-simultaneous” and “simultaneous” judgments, respectively (cf. García-Pérez & Alcalá-Quintana, 2012), but λ can also include task-specific constraints that prevent perfect performance even in the absence of errors (see van Eijk et al., 2010). The four-parameter Gaussian model was fitted to the data using again the R package *quickpsy* (Linares & López-Moliner, 2016), by exploiting quickpsy's option to include a user-defined function. Absolute model fit was again assessed using R^2. Using the estimated psychometric functions, the PSS and an index of the spread of the function were computed for the function Φ, i.e., after the error parameters had been removed. Different from the (typical) shape of the complete psychometric function, Φ ranges from 0 to 1.

The PSS for simultaneity judgments is the point on the x-axis (SOA) at which the Gaussian function reaches its peak, and corresponds to the function's μ parameter (see Figure 2). As an index of the spread of the Gaussian function, we used the half-width half-height (HWHH; Fujisaki & Nishida, 2009). This is half the width of the Gaussian function at half its height; for the error-free function, it is half of the width of the function at p = .50. An advantage of the HWHH is that, by referring to the standard normal distribution, it can be converted into an equivalent of the JND for TOJs (JND = HWHH/1.746), and thus can be directly compared to, and meta-analytically integrated with, the JND obtained in the TOJ task (Fujisaki & Nishida, 2009).

Results

Preliminary data analyses revealed that four participants did not meet our inclusion criteria, meaning that the final sample comprised 16 participants. One of the excluded participants gave zero or very low unpleasantness ratings to most pictures and three had anomalous simultaneity judgments for affect. In one case, the estimated peak of the curve fell outside the SOA range. In the second case, the probability of “simultaneous” judgments remained below .50 at all SOAs; in the third case, affect was judged as simultaneous to the perception of the probe at all but one SOA. The response patterns of these two participants suggest that they found the task too difficult and resorted to, respectively, random or constant responding (see e.g., Kostaki & Vataki, 2018; Love et al., 2011; Spence et al., 2003). However, as reported in Supplementary Materials S1, the mean PSSs and JNDs changed only little if the excluded participants were kept in the analyses. In three of the four excluded cases, the PSSs for OR and AF were in line with predictions (OR < AF), but the correlation between OR and AF was reduced from r =.31 to r = -.03 if all participants were included.

Table 3
Fit Index (R^2), PSS and JND for the SJ Task in Experiment 2

	R^2			PSS			JND		
Task	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Object recognition	.83	.18	[.38, .99]	11	24	[-14, 64]	53	22	[19, 96]
Affect	.88	.08	[.68, .99]	165	149	[9, 512]	152	87	[45, 393]

Note: $N = 16$.

Affect Ratings

The analysis of the affect ratings yielded findings similar to those obtained in Experiments 1a and 1b. In particular, the pictures (a) induced at least moderately intense pleasant or unpleasant affect in all included participants and (b) largely retained their affect-inducing capacity from the pre-SJ rating (pleasant: $M = 8.1$, unpleasant: $M = 6.4$) to the post-SJ rating (pleasant: $M = 7.7$, unpleasant: $M = 5.1$).

Simultaneity Judgments

The average statistical fit, PSS and observer sensitivity (the JND computed from the HWHH) of the 16 included participants are shown in Table 3. Figure 2 shows the psychometric function for the averaged parameters of the individual participants after removal of the error parameters. The curves of all 20 individual participants can be found in Supplementary Materials S1 (Figure S1.3).

Model Fits. The average fit of the Gaussian function was $R^2 = .83$ for OR and .88 for AF (Table 3). Although the model fits were lower than those obtained for TOJs in Experiments 1a and 1b, inspection of the individual psychometric functions (see Figure S1.3 in Supplementary Materials S1) suggests that the main reason for this was that the simultaneity judgments contained a higher amount of random error. Nonetheless, the peak of the function could be well discerned in all included cases.

PSSs for OR and AF. As in Experiments 1a and 1b, the ANOVA conducted on the individual PSSs revealed a significant effect of task, $F(1, 12) = 14.32$, $p = .003$, $\eta_g^2 = .36$, reflecting that the participants became aware of the affect significantly later (165 ms) than they recognized the objects (11 ms). No other main effect or interaction reached significance, $F\text{s} \leq 1.52$, $p\text{s} \geq .981$. Thus, the participants experienced the affect, on average, 154 ms after they recognized the object. A positive time lag was obtained for all 16 participants. Thus, Hypothesis 1 (OR < AF) was again supported.

Correlations Between the Latencies of OR and AF. The correlation between the latencies of OR and AF was again positive but not significant given the sample size, $r(14) = .31$, $p = .244$ (two-sided test). Hence, Hypothesis 2 again received only limited support.

JNDs. The ANOVA for the JNDs (computed from the HWHH) yielded a significant effect of task type, $F(1, 12) = 23.68$, $p = .003$, $\eta_g^2 = .45$, reflecting a larger spread of the psychometric function for AF (152 ms) than OR (53 ms). None of the other main or interaction effects was significant, $F\text{s} \leq 1.36$, $p\text{s} \geq .770$.

Replication With the GPAQ-SJ Model. To check the robustness of the PSS and JND estimates across estimation methods, we fitted the GPAQ-SJ2 model (García-Pérez and Alcalá-Quintana, 2012) to the data. Details are reported in Supplementary Materials S1. Analogous to the TOJs (Experiment 1a), the GPAQ-SJ2 model fit the data significantly better than the Gaussian function model. Again, however, the average PSS (OR: 17 ms, AF: 159 ms) and JND (OR: 47 ms, AF: 144 ms) estimates were very similar to those obtained with the Gaussian function model. The correlation between the individual PSS estimates obtained with the two estimation methods was again high (r = .98) for AF, although only moderate (r = .56) for OR.

Analysis for Pictures as Units. Due most likely to the low number of SJ data per picture (9 for the pleasant, 7 for the unpleasant pictures), the by-pictures analysis yielded somewhat lower fits than the by-subjects analysis (M = .82, SD = .09 for OR and M = .81, SD = .10 for AF). Nonetheless, the average PSS for OR (8 ms, SD = 15 ms) and AF (153 ms, SD = 72 ms) was again very similar to those obtained in the by-subjects analysis. An ANOVA with task and picture valence confirmed the main effect of task, $F(1, 22) = 101.44$, $p < .001$, $\eta_g^2 = .68$. Neither picture valence nor the interaction was significant, Fs ≤ 1.40, ps $\geq .500$. The PSS of AF occurred later than the PSS of OR for all 24 pictures (M = 145 ms, SD = 69 ms). The JND for AF (163 ms) was again significantly higher than that for OR (M = 47 ms), $F(1, 22) = 47.54$, $p < .001$, $\eta_g^2 = .53$.

Discussion

Experiment 2 replicated the main findings of Experiment 1a and 1b using an SJ task instead of a TOJ task. As in the previous experiments, the participants experienced the affect evoked by the pictures on average significantly later (in Experiment 2, by 154 ms) than they recognized the depicted object; this effect was not significantly influenced by task order (OR or AF) and picture valence (pleasant or unpleasant), and it was also obtained on the individual level (in Exp. 2, for all participants). Also replicating Experiments 1a and 1b, the by-pictures analysis confirmed the later PSS for AF than OR; in Experiment 2, this effect was obtained for all pictures.

The main difference in the results compared to those of Experiments 1a and 1b was that the time lag between OR and AF was greater in the SJ experiment (154 ms) than in the first and second TOJ experiment (79 ms and 113 ms, respectively). However, part of this difference was due to the earlier PSS of OR obtained in the SJ experiment. Because the latter PSS estimate was obtained with a more precise (higher-resolution) measurement procedure (more narrowly spaced SOAs), it can probably be regarded as the more precise estimate. Using this estimate (11 ms), the time lag between OR and AF in Experiment 1a increases to 110 ms and in Experiment 1b, to 123 ms, which approaches the delay estimated with the SJ task in Experiment 2.

Experiment 3: Generalizing the Findings

Although the results of Experiments 1a, 1b and 2 support the semantic primacy hypothesis, the studies have two limitations. First and most important, the sample of affective pictures was limited, because only 12 pleasant and 12 unpleasant pictures were presented, and the unpleasant pictures were all of the disgust type (with fear as a possible secondary affective reaction to some stimuli, e.g., pictures of snakes and a spider). Second, the pleasant and unpleasant pictures were judged in separate blocks of trials. As mentioned, this was done to facilitate the judgment of affect onset, by allowing the participants to focus on the onset of a feeling of known valence. However, it could be argued that this procedure resulted in an accumulation of picture-evoked feelings of pleasure or displeasure across trials and that, as a consequence, it became increasingly difficult for the participants to detect the onset of affect evoked by a picture against the rising baseline. If so, this could have resulted in an artifactually prolonged latency of affect detection.[5]

To overcome these potential limitations of the preceding studies, another TOJ experiment was conducted. Addressing the first limitation, a much larger number of pictures (84 pleasant and 84 unpleasant) was presented. Addressing the second issue, pleasant and unpleasant pictures were presented randomly intermixed in the TOJ tasks. As a consequence, pleasant pictures were regularly followed by unpleasant pictures and vice versa, which should have prevented an accumulation of same-valence affect across trials, if it indeed occurred in the previous experiments.

In addition, each picture was only presented once in each TOJ task, which allowed to test whether the findings of Experiments 1a, 1b and 2 generalize to comparatively novel pictures.

Method

Participants

We recruited 40 participants, 33 females and seven males (M age = 25.7, SD = 8.0). Twenty-four were students (17 psychology) and one attended high-school; the other participants were employed. Because of the increased difficulty of recruiting participants for laboratory studies during the Sars-CoV-2 epidemic in 2021, we also created an online version of the experiment. Sixteen participants took part in the laboratory and 24 in the online experiment. Most psychology students took part as part of their study requirements, but some of them received a monetary compensation, as did the rest of the participants. Data of four participants of the online experiment were incomplete, most likely because of internet connection problems, and therefore had to be discarded. Two additional participants were excluded from the data analyses because they had very bad curve fits, suggesting near-random responding (see Results), reducing the effective sample size to $N = 34$ (16 lab, 18 online). Results were virtually unchanged if these two participants were included (see Supplementary

[5]We did not consider this to be a serious possibility in Experiment 1–2 because we assumed that the picture-evoked affect would terminate quickly after picture offset, particularly because the pictures were followed by the answer screen (see Figure 1), which should have worked as a mask (i.e., should have terminated further picture processing). In any case, it seemed unlikely to us that any residual pleasure or displeasure would survive the fairly long interval between the offset of one affective picture and the onset of the next (about 1100–2300 ms depending on whether the probe appeared before or after the target picture and the SOA, and given the average response time of ~600 ms). Indirect support for this assumption was obtained from trend analyses of the affect ratings for pleasant and unpleasant pictures across trials in experiments 1a, 1b and 2, using a random-intercept linear mixed model. This analysis revealed no evidence for an accumulation of affect across trials in the rating task, with b for trial number ranging from -0.08 (suggesting a slight decrease of affect) to 0.02, $ps > .142$.

Materials S1).

Design

Different from Experiments 1a, 1b and 2, 168 pictures, 84 pleasant and 84 unpleasant, were shown in the OR and AF TOJ tasks. Task order (OR or AF TOJs first) was counterbalanced across subjects to control for possible sequence effects. Participants were randomly assigned to the two task order groups.

Procedure

The laboratory version of the experiment was programmed in DMDX. The web version was programmed in jsPsych (de Leeuw, 2014) and mimicked the lab version as closely as possible. The key assignment used in the lab experiment could not be used in the Web version because it was Windows-specific. Therefore, the participants were instructed to use the keys “d” and “f” (adjacent keys on the German keyboard) to enter -2 and -1 with the middle and index finger of their left hand, and “j” and “k” (also adjacent keys) to enter +1 and +2 with the index and middle finger of their right hand.

Experiment 3 comprised five consecutive tasks: an abbreviated version of the affect rating task (see below), three TOJ tasks (color recognition, OR, and AF), and a repetition of the abbreviated affect rating task, plus ratings of the task difficulty of the TOJ tasks. All tasks were completed in a single session, either in a computer lab of the Institute of Psychology (with one subject participating at a time due to COVID-19 restrictions), or online with subjects participating from home. The experiment took 30 to 45 minutes.

Affective Stimuli. Eighty-four pleasant and 84 unpleasant pictures from the OASIS picture pool of 900 affective stimuli (Kurdi et al., 2016) were selected in a three-step procedure. First, the 144 most pleasant and the 144 pictures most unpleasant pictures according to the normative data provided by Kurdi et al. (2016), were selected from the OASIS picture set. Second, these pictures were rated by 20 students in an online study programmed with jsPsych (de Leeuw, 2014) for the intensity of pleasure or displeasure they evoked, on a bipolar rating scale ranging from -7 (evokes an extremely unpleasant feeling) to 0 (evokes no feeling) to +7 (evokes an extremely pleasant feeling), and the 84 pleasant and 84 unpleasant pictures with the highest average ratings were selected. Third, pictures of objects that were overrepresented or difficult to recognize were replaced by, respectively, pictures of underrepresented object categories or by better-recognizable pictures from the same object category (accepting that the replacement pictures had slightly lower affect ratings). For example, pictures of lakes were initially overrepresented among the pleasant pictures; therefore, some of them were replaced by pictures of fireworks, flowers and desserts. The final picture set had average ratings of evoked pleasure from 2.0 to 4.4 (M = 3.1, SD = 0.5) and of evoked displeasure from -1.9 to -4.3 (M = -2.71, SD = 0.5). The pictures were presented in their original size (500 × 400 pixels). The OASIS picture names can be found in Supplementary Material S1.

Affect Rating Task. To keep the experiment as short as possible, the 168 pictures were not individually rated in Experiment 3. Instead, they were presented as a picture stream (1240 ms per picture) in an individually randomized order and the participants subsequently estimated the percentage of the pictures (a) where they recognized the depicted object or objects and (b) that evoked at least a weak pleasant or unpleasant feeling. Answers were given on a rating scale with 11 categories, ranging from "less than or equal to 50%" to "100%", with 5% steps in between.

TOJ Tasks. The TOJ tasks were based on those used in Experiment 1a, but two SOAs (-1200 ms and +1200 ms) were added to allow all participants to reach the asymptotes of the psychometric curve and thus, improve curve fitting (see Experiment 1b). As in the previous experiments, 12 pictures were presented at each SOA. In Experiment 3, however, pleasant and unpleasant pictures were intermixed, and each picture was presented only once. To counterbalance possible effects of picture valence, six pleasant and six unpleasant pictures were presented at each SOA. The identity of the pleasant and unpleasant pictures assigned to the 12 SOAs was randomly determined for each participant; however, to maximize the comparability of the OR and AF tasks, the same picture-SOA assignment was used in both tasks.

Post-Experimental Ratings. Following the TOJ tasks, the participants (a) again estimated the percentage of recognized and affect-evoking pictures during the task; (b) indicated which key or keys they had used if the object was not recognized or no affect was experienced; (c) estimated, for both the OR and AF task, the percentage of trials in which they had experienced uncertainty about their answers on an 11-point-scale ranging from 0 to 100% in steps of 10%; and (d) rated the subjective difficulty of both tasks on an 11-point-scale ranging from 0 (not difficult at all) to 10 (extremely difficult). These ratings were collected to explore possible reasons of bad psychometric curve fits or unpredicted PSS differences.

Results

Affect and Post-Experimental Ratings

The details of the analysis of the affect estimates and the post-experimental ratings can be found in Supplementary Material S1. Briefly, these findings were: (a) According to the participants' frequency estimates, nearly all pictures could be recognized and about 80% evoked a pleasant or unpleasant feeling in the majority of the participants; (b) the OR task was on average experienced as easy (M = 2.6, SD = 2.0 on the 0–10 scale) and the participants felt certain about their answers in about 80% of the trials, whereas the AF task was experienced as being of medium difficulty (M = 4.8, SD = 2.6) and the participants were uncertain about their answers in about 70% of the trials. However, most participants still experienced the AF task as feasible, and the uncertainty and difficulty ratings of AF were essentially uncorrelated with the PSS of AF (rs = .01 and -.04, respectively).

Temporal Order Judgments

Table 4 summarizes the individual fits (R^2) of the probit model to the OR and AF TOJs, as well as the PSSs and JNDs computed on the basis of the latent psychometric functions (i.e., after removing the error parameters γ and λ). Two participants had bad fits (R^2 = .42 and .52) of the psychometric function for affect and one of them (with R^2 = .52) did not exceed 50% within the SOA range. The post-experimental ratings revealed that these participants judged the affect task as very difficult (M = 8.5) and felt uncertain about their answers in most trials (M = 75%). We therefore removed these participants from the analyses reported below, but as reported in

Table 4
Fit Index (R^2), PSS and JND for the TOJ Task in Experiment 3

	R^2			PSS			JND		
Task	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Object recognition	.97	.03	[.87, >.99]	61	103	[-44, 404]	63	61	[2, 296]
Affect	.95	.04	[.83, >.99]	139	184	[-72, 786]	122	91	[7, 345]

Note: $N = 34$.

Supplementary Materials S1, the mean results were essentially unchanged if these participants were included. However, the Pearson correlation between OR and AF was reduced from $r = .67$ to $r = .38$.

Model Fits. The average fit of the probit model to the TOJs was on average $R^2 = .97$ ($SD = .03$) for OR, and $R^2 = .95$ ($SD = .04$) for AF. To check whether there was a difference between the web and laboratory versions of the experiment, we conducted a mixed-design ANOVA on the fit index, with experiment version (lab or web) and task order (AF first or second) as between-subject factors, and task (OR or AF) as within-subject factor. The *p*-values for experiment version and task order were Holm-corrected. This ANOVA revealed no significant main effects or interactions, Fs ≤ 6.29, ps $\geq .125$.

PSSs of OR and AF. A 2 (task) × 2 (task order) × 2 (experiment version) ANOVA of the individual PSSs revealed a significant effect of task, $F(1, 30) = 9.63$, $p = .004$, $\eta^2_g = .06$. No other effect reached significance, Fs ≤ 2.42, ps $\geq .781$. The mean PSS was 61 ms for OR and 139 ms for AF (Table 4); hence, the participants became on average aware of the picture-evoked feelings 78 ms after they recognized the object, again supporting Hypothesis 1.

Twenty-six of the 34 participants conformed to the mean pattern (AF > OR); the remaining nine showed the opposite pattern, with a mean PSS difference of -58 ms ($SD = 27$ ms). These participants judged both the OR and AF task as more difficult (Ms = 3.5 and 5.6, respectively) than the rest (Ms = 2.3 and 4.5, respectively).

Correlation Between the Latencies of OR and AF. The correlation between the latencies of OR and AF in Experiment 3 was $r(32) = .67$, $p < .001$ (two-sided test). Hence, in Experiment 3, Hypothesis 2 was clearly supported.

JNDs. As in the previous experiments, the spread of the psychometric function (JND) was greater for AF ($M = 122$ ms) than for OR ($M = 63$ ms); and as in Experiment 2, this difference was significant, $F(1, 30) = 12.89$, $p = .010$, $\eta^2_g = .12$. None of the other effects was significant, Fs ≤ 3.45, ps $\geq .438$.

A by-pictures analysis was not possible in Experiment 3, because each picture was only presented once in each TOJ task.

Discussion

Experiment 3 found that the model fits, PSSs, and JNDs of the previous TOJ experiments (Experiments 1a and 1b) generalize well to a larger picture set and to an experimental condition where each picture is presented only once, and where pleasant and unpleasant pictures are intermixed. As before, the participants became aware of the feelings evoked by the affective pictures significantly later than they recognized the object shown in the pictures, with an average delay of 78 ms. Again, the cognition-affect delay was independent of the order of the OR and AF TOJ tasks; hence, there was no indication of a training or reminiscence effect. In addition, no significant difference was found between the laboratory and online version of the experiment.

Again, the predicted cognition-affect lag was again found for most participants. Possible reasons for deviating mean patterns (OR > AF) are discussed in the General Discussion. The JND for affect was found to be significantly greater than that for object recognition, replicating Experiment 2.

The correlation between the latencies of OR and AF (Hypothesis 2) was much larger in Experiment 3 and statistically significant. The reason may have been that the between-subjects variance of the latencies was larger in Experiment 3 than in the preceding experiments (see Tables 1–4). The higher variances of the latencies are plausible because the pictures were less familiar in Experiment 3.

Meta-Analytic Integration of the Findings and the Effects of Affect Intensity

To obtain cross-experiment and cross-method estimates of the size of the cognition-affect lag and the cognition-affect correlation and to profit from the higher precision of the combined estimates and the higher power of the statistical tests (see Schimmack, 2012), the results of the four experiments were meta-analytically integrated. In addition, we conducted an exploratory analysis to investigate the possible effect of affect intensity on the latency of affect.

PSS and JND

Figure 3 graphically displays the mean estimates of the PSS and JND for OR and AF obtained in the four experiments plus the meta-analytically integrated performance measures, together with their confidence intervals. The fixed-effects meta-analysis model was used because we assumed that the latency measurements of OR and AF in the different experiments estimate the same perceptual latencies (see Borenstein et al., 2009). The meta-analytic PSS mean (computed using the *metamean* function in the R package *meta*; Balduzzi et al., 2019) was 19 ms for OR and 133 ms for AF (delay = 114 ms) in the by-subjects analysis. Nearly the same values were obtained in the by-pictures analysis: 16 ms for OR and 134 ms for OR (delay = 118 ms). The combined difference between the experiment-wise latencies is highly significant in both the by-subjects and the by-pictures analysis, $p < 10^{-8}$.

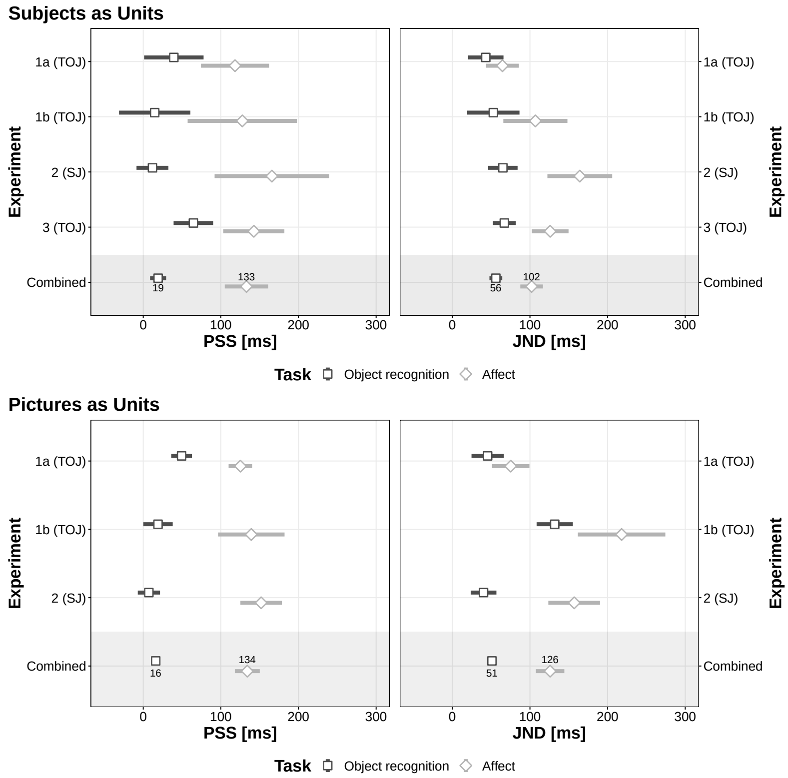
Note that Experiment 3 could not be analyzed by-pictures because each picture was only judged once in each task.

Correlations Between the Latencies of Object Recognition and Affect

Table 5 shows the Pearson correlations between the two PSSs for the subject- and picture-centered analyses in the four experiments, as well as the meta-analytically

Figure 3

Average PSS and JNDs Obtained in Experiments 1–3 and Meta-Analytically Combined Performance Measure



Note: PSSs (left panels) and JNDs (right panels) for object recognition and affect obtained in the four experiments and the meta-analytically combined performance measures together with the respective 95%-CIs, for subjects as units (upper panel) and pictures as units (lower panel). A by-pictures analysis was not possible in Experiment 3 because each picture was only presented once in each task. The CIs of the uncombined performance measures are the 95%-Cousineau-Morey-CIs (Baguley, 2012) of the means.

combined r (we used the weighted average after a Fisher r-to-z transformation). We report the p-values for a two-tailed z-test on the aggregated correlations.

As summarized in Table 5, positive correlations were obtained in all four by-subjects analyses and in two of the three possible by-pictures analyses (no by-pictures analysis was possible in Experiment 3). For the by-subjects analyses, the combined correlation was r = .51, 95%-CI = [.27, .74], p < .001. For the by-pictures analysis, the combined correlation was r = .13, 95%-CI = [-.12, .38], p = .300.

To evaluate how much the correlations were attenuated by measurement error, we estimated the reliability of the by-subjects PSS estimates obtained in the four experiments using the following parametric bootstrap method: (1) For each participant, 100 bootstrap samples of the TOJs or SJs were generated and the psychometric models were re-estimated from these samples; (2) the PSS estimates obtained from fitting the 1st, 2nd etc. bootstrap sample of each participant were correlated with each other; (3) the median of the correlations was computed after a Fisher r-to-z transformation. This estimation method yielded reliabilities ranging from .82 to .85 for the OR latencies and of .74 to .90 for the AF latencies; for experiment-wise

Table 5

Correlations Between the PSSs of Object Recognition and Affect, and Correlations of the PSS of Affect with Affect Intensity, for Subjects and Pictures as the Units of Analysis

	Subjects as Units				Pictures as Units	
Experiment	OR-AF		AI-AF		OR-AF	AI-AF
1a (TOJ)	.26	(.43)	.15	(.20)	.24	-.15
1b (TOJ)	.19	(.29)	.32	(.42)	-.15	.02
2 (SJ)	.31	(.41)	-.13	(-.14)	.29	-.05
3 (TOJ)	.67	(.88)	-		-	-
Combined	.51	(.85)	.10	(.15)	.13	-.06

Note: OR = Object recognition latency (PSS), AF = Affect latency (PSS), AI = Affect intensity. The correlations in parentheses (in the by-subjects analyses) have been corrected for attenuation based on the reliabilities of the latencies reported in Table 6.

Table 6

Reliabilities of the PSSs of Object Recognition and Affect for Subjects as the Units of Analysis

Experiment	**Object Recognition**	**Affect**
1a (TOJ)	.82	.74
1b (TOJ)	.84	.77
2 (SJ)	.85	.90
3 (TOJ)	.85	.89

reliabilities, see Table 6. Corrected for attenuation, the correlation between the PSS for OR and AF (for subjects as units) increases to .43 (Exp. 1a), .29 (Exp. 1b), .41 (Exp. 2), and .88 (Exp. 3) and the meta-analytically combined correlation is r = .85.

Effect of Affect Intensity on Affect Latency

Stimulus intensity has been found to speed up the perceptual latencies in sensory temporal judgment tasks (Jaśkowski, 2014; Neumann & Niepel, 2004): Stimuli of high intensity are perceived earlier. Because affective feelings also have an intensity, and are in fact often likened to sensations (see Reisenzein, 2012), a parallel effect could be expected for affect intensity. The data of experiments 1a, 1b and 2 allow to test this hypothesis, if only for moderate to high affect intensities (because the stimuli were preselected to evoke at least moderate affect), by correlating the individual PSSs for affect with the mean ratings of affect intensity for participants and pictures. The hypothesis would be supported by a significant negative correlation between affect intensity and affect latency. Note that this analysis was exploratory.

With the exception of the by-pictures analysis in Exp. 1b (that, however, was positive rather than negative), the obtained correlations were very low and not significant (meta-analytic r = .10, 95%-CI = [-.22, .42], p = .529, for subjects, and r = -.06, 95%-CI = [-.31, .19], p = .621, for pictures). Hence, no evidence for the affect intensity→affect latency hypothesis was found. However, because the average affect intensities for pictures and participants were fairly high in experiments 1–2 (> 5 on the 0–10 scale), it is possible that an intensity effect on affect latencies will be found if affective stimuli of lower intensity are included. With respect to the present studies, however, the lack of correlation between affect intensity and affect latency suggests that the obtained delay between OR and AF cannot be attributed to the inclusion of a few low-intensity affect pictures for which affect onset was hard to detect. This was confirmed by the results of an additional exploratory analysis of the data of Experiments 1a, 1b and 2, which found nearly identical PSSs if the participants' psychometric functions were re-estimated after pictures rated < 5 on the affect intensity scale by the participant had been excluded (details are available from the authors).

General Discussion

The studies reported in this article used, for the first time, temporal order and simultaneity judgments to estimate both the subjective time of occurrence of object recognition and emotional feelings (pleasure and displeasure). The main findings can be summarized as follows.

First, supporting Hypothesis 1, in both kinds of tasks the feelings elicited by the pictures were on average noticed significantly later than the depicted objects were recognized. If we take the meta-analytically combined means as the best estimates of the PSSs of object recognition and affect, the cognition-affect lag was on average 114 ms; if the probably more accurate PSS of object recognition obtained in the SJ study (11 ms) is used, the lag increases to 122 ms. For pictures as the units of analysis (Exp.1a, 1b, and 2), nearly the same delay was found (118 ms). In the rest of the discussion we focus on the result of the by-subjects analysis and use ∼120 ms as the estimate of the size of the cognition-affect lag.

Second, the direction and size of the temporal lag between object recognition and affect was independent of task order (object recognition vs. affect judgment first) and the hedonic tone of the affect evoked by the pictures (pleasure vs. displeasure). For the first TOJ experiment (Exp. 1) and the SJ experiment (Exp. 2), we additionally checked whether the PSS estimates are robust across different psychometric estimation methods, by also fitting the psychometric models for TOJs and SJs proposed by García-Pérez and Alcalá-Quintana (2012) to the data. Evidence for robustness was obtained. Furthermore, the direction of the temporal lag was largely independent of the specific content of the pictures, as it was found for all (Exp. 2) or nearly all (Exp. 1a and 1b) pictures; however, the size of the delay varied considerably for different pictures.

Third, the direction of the temporal difference obtained for the mean PSS was also found on the individual level for most participants (69 of 81). However, as found for pictures, substantial differences in the size of the temporal delay were also found for different participants.

Fourth, the PSSs of object recognition and affect were, as predicted, positively correlated in all experiments in the by-subjects analyses, and in two of the three studies where the by-pictures analysis could be computed. For participants but not pictures, the correlation reached significance in Experiment 3, as well as in the meta-analysis. The by-subjects correlations ranged from .19 to .67; the meta-analytic average was .51.

Fifth, Experiment 3 found that the results generalize well to a much larger set of pleasant and unpleasant pictures, that are presented in intermixed sequence, and are comparatively novel to the participants because they were only repeated once in each task.

Sixth, the spread (indicated by the JND) of the participants' psychometric functions for affect was consistently greater than that for object recognition, and this difference was statistically significant in experiments 2 and 3 as well as in the meta-analysis. This finding suggests that the latencies of affect onset are more variable than those of object recognition, a conclusion that was supported by the by-pictures analyses. This finding provides additional support for the hypothesis that affect generation not only follows, but depends on, object recognition. As noted by Nummenmaa et al. (2010), this hypothesis suggests that the latency of affect is the sum of the processing times for object recognition and affect generation. If so, the variance of the affect latencies must be greater than that of object recognition (e.g., Sternberg, 1969).

Possible Reasons for Deviating Response Patterns

The finding that 12 of the 81 included participants had a shorter latency for AF than OR merits discussion. One possible explanation is that these participants belong to a minority in which affect is indeed evoked pre-conceptually, as the affective primacy hypothesis holds. Our data do not allow to decide this hypothesis, but it could be tested in future studies, for example by examining whether the negative cognition-affect delay of these participants is stable across time and different latency measurement methods (e.g., TOJ/SJ, sRT). However, in our view, a more plausible explanation is that the unpredicted latency differences reflect measurement problems. In particular, we suspect that these participants found the affect TOJ/SJ too difficult and therefore replaced it, deliberately or unwittingly, by a simpler judgment strategy.

The finding that some participants find TOJ or SJ tasks too difficult is not unusual (e.g., Love et al., 2013). The assumption that such participants were also present in our studies is supported by (a) the response patterns of most of the excluded participants, whose data were either very noisy, indicating near-random responding (Exp. 2 & 3) or who gave constant responses at nearly all SOAs (one participant in Exp. 2); and (b) the finding of Experiment 3 that the participants with deviating response patterns rated both the object recognition and affect task to be more difficult than the remaining participants. Our hypothesis is that the participants with regular psychometric curves but deviating latency differences also found the affect task too difficult; but instead of resorting to random responding, they used a simplifying judgment strategy. Specifically, instead of judging the time of affect onset, they may have judged the time of picture onset, or the time of object recognition. In the first case, the PSSs of object recognition and affect would be reversed; in the second case, the two PSSs would be identical up to random error. A third possibility is that these participants judged when they noticed that the picture was pleasant or unpleasant (i.e., when they detected its valence), rather than when they felt the affect elicited by it, and retrieved the valence of the pictures from memory. In this case, the PSS for affect should again be close to the PSS for object recognition. Evidence for this latter hypothesis was obtained in an sRT study by Franikowski and Reisenzein (in press), who found that valence judgments of affective pictures are on average faster than judgments of affect onset, and that their latency correlates with the latencies of both affect and object recognition.

Comparison to Previous Findings

The central finding of our studies, that affect sets in with a time lag after object recognition, is consistent with the results of previous studies, reviewed in the introduction, that investigated the same question using sRT tasks (Nakashima, 1909a; 1909b; Nummenmaa et al., 2010). Most of our other findings, summarized above, also agree with parallel findings reported by Nummenmaa et al. (2010): Stimulus valence and task order had no effect on the delay between cognition and affect; the delay was found for most participants and most pictures, while at the same time, there were noteworthy differences between participants and between stimuli; the estimated onset times of object recognition and affect were moderately positively correlated (Nummenmaa et al., Exp. 3., reported a correlation of .43 for the by-pictures analysis); and the spread of the distribution of the RT latencies for affect was broader than that of the latencies for object recognition (see e.g., Fig. 5 in Nummenmaa et al., 2010, p. 231).

These agreements between our results and those of the sRT studies are noteworthy not only because they were found despite the different time measurement methods used, but also because of several other differences between studies. These concern in particular the presentation times of the stimuli in our studies (up to 3307 ms in Exp. 1b versus 30 ms in the studies of Nummenmaa et al., 2010) and the wording of the cognitive and affective judgments required of the participants. One or both of these differences could explain the main remaining difference between our studies and those of Nummenmaa et al., (2010), the finding that the cognition-affect delay obtained in our studies (∼120 ms) was noticeably *larger* than the delay found in the sRT studies of Nummenmaa et al. (2010, about 40–80 ms), despite the fact that differences in latencies estimated with temporal judgments are typically 2–4 times *smaller* than if measured with sRTs (Cardoso-Leite & Gorea, 2010).[6]

Partly to address this issue, we recently conducted two sRT experiments (Franikowski & Reisenzein, in press) with the same (Exp. 1) or largely the same (Exp. 2) pictures as in the present experiments 1–2, using similar presentation times and the same OR and AF tasks, plus a valence detection task in which the participants had to decide as quickly as possible whether the picture was pleasant or unpleasant (Fazio, 1995). In agreement with previous findings comparing sRT and TOJ/SJ tasks (Cardoso-Leite et al., 2010), as well as with Nakashima's (1909a; 1909b) sRT experiments, our sRT studies found that the object recognition and affect latencies were much longer than in the present TOJ/SJ studies. More important, the delay between object recognition and affect was significantly longer (∼500 ms) than the delay between object recognition and valence detection (∼170 ms), and the latter delay was still about three times larger than in the sRT studies of Nummenmaa et al. (2010). These findings suggest that the differences in the size of the cognition-affect lag found in

[6] A parsimonious explanation of this difference is that participants use a higher criterion for deciding on the presence of a signal in sRT tasks than in temporal judgment tasks (see Cardoso-Leite & Gorea, 2010; Miller & Schwarz, 2006).

the present studies and those of Nummenmaa et al. (2010) can be mainly attributed to the fact that the participants performed different tasks (the detection of affect onset detection versus of stimulus valence); whereas the longer delay between OR and valence detection in the sRT studies of Franikowski and Reisenzein (in press), compared to Nummenmaa et al. (2010), can be explained by the longer picture presentation times used in the former studies: The longer presentation times may have encouraged the participants to delay the response to increase their certainty that they had detected the valence of the stimulus. In contrast, with a stimulus presentation time of 30 ms (Nummenmaa et al., 2010) there is nothing to be gained from waiting, as no additional information about the stimulus is forthcoming.

Implications of the Findings for Emotion Research

Our findings have implications for several areas of emotion research. Here, we focus on the implications of the findings for the cognition-emotion debate and for bodily feedback theories of emotion.

Cognitive vs. Noncognitive Theories of Emotion Generation

Our findings support the semantic primacy hypothesis of emotion generation, that emotions presuppose the semantic categorization of evoking objects. In fact, the data are consistent with the stronger hypothesis that the experience of pleasant and unpleasant feelings evoked by pictures is based on the *conscious* recognition of the depicted objects. The findings thus add to the evidence from recent subliminal perception studies (Lähteenmäki et al., 2015; Peira et al., 2012; see also, Hedges et al., 2015) and the evidence from sRT studies (Nakashima, 1909a; 1909b; Nummenmaa et al., 2010) which support the hypothesis of semantic primacy. By the same reasoning, our findings speak against the hypothesis that the affect evoked by pictures is caused by subconscious processes that precede recognition of the depicted objects (e.g., LeDoux, 1998; Murphy & Zajonc, 1993; Öhman & Mineka, 2001).

Bodily Feedback Theories of Emotional Experience

A second long-standing question of emotion psychology for which our findings are relevant concerns the role of bodily feedback for emotional experience (see Reisenzein & Stephan, 2014). Beginning with James (1890), many emotion theorists have proposed that bodily feedback, in particular feedback from peripheral physiological reactions, is necessary for emotional experience, or at least adds importantly to its quality and/or intensity (e.g., Damasio, 1984; Laird, 2007; Schachter, 1964). Even today, this hypothesis is widespread (Reisenzein & Stephan, 2014). However, this hypothesis faces several objections (Reisenzein & Stephan, 2014). Of greatest relevance in the present context is the objection, first raised by Stumpf (1899) and subsequently stated more forcefully by Cannon (1927), that the latency of the physiological reactions evoked by affective stimuli is too long for them to be the basis of emotional experience. Defenders of bodily feedback theory have responded that some physiological reactions evoked by emotional stimuli can occur more quickly than Cannon (1927) had claimed, and that feedback from facial expressions (not considered by Cannon) is even faster (e.g., Laird, 2007). For example, the latency of heart rate changes evoked by pleasant and unpleasant pictures of the kind used in our experiments can be as short as one second (e.g., Bradley et al., 2001) and the activation of facial muscles (zygomaticus and corrugator) can begin as early as 500 ms after stimulus onset (e.g., Dimberg et al., 1998).

However, these latencies – at least those of physiological reactions – seem still too long to fit comfortably into the time window allowed by the latency of affect suggested by the early sRT studies (Nakashima, 1909a, 1909b; Wells, 1925), particularly because (a) additional time is needed for feedback from the periphery to reach the brain, and (b) the RT for affect also includes the time needed to select and execute the motor response. Certainly, the latencies of bodily reactions do not fit within the time window of about 350 ms obtained by Nummenmaa et al. (2010, Exp. 1–3) for valence judgments of pictures.

The findings of our studies provide further support for this conclusion. Although, as mentioned in the introduction, the latency of affect estimated with temporal judgment tasks is only a relative time estimate (the time difference to the detection of the probe), it can be used to estimate the absolute latency of affect if the absolute latency of either probe detection or object recognition is known. Good estimates of the latency of conscious object recognition are available from the studies of Kirchner and Thorpe (2006; see also Fabre-Thorpe, 2011) and Nummenmaa et al. (2010, Exp. 2–3), who used eye-gazes, which are much faster than manual responses, as the response mode. Kirchner and Thorpe (2006) reported an average median RT latency for object recognition (semantic classification) of about 230 ms, whereas Nummenmaa et al. (2010) reported a median RT of about 290 ms. Referring to findings by Schiller and Kendall (2004), Kirchner and Thorpe (2006) assume that the time for preparing a saccade is about 30 ms; hence the average latency of conscious object recognition is in the range of 200–260 ms. This estimate agrees fairly well with the findings of ERP studies of object recognition (see e.g., Fabre-Thorpe, 2011; Johnson & Olshausen, 2003). Adding the average delay of ∼120 ms for affect obtained in our studies, one can conclude that picture-evoked pleasant and unpleasant affect is felt, on average, between 320–380 ms after picture onset. The lower bound of this interval is nearly identical to the latency of the potentiation of the startle reaction by aversive pictures found by Globisch et al. (1999) and the latency of the affective modulation of the LPP (late positive potential) component of the ERP (Hajcak & Foti, 2020). However, even assuming that 380 ms is the more accurate estimate, the latency of affect seems too short even for fast facial feedback (Dimberg et al., 1998) to be its basis. Certainly, the latency of 380 ms is too short for feedback from peripheral physiological reactions, whatever their nature, to ground affect.

It could be argued that, even if the bodily feedback theory cannot account for pleasant and unpleasant feelings, it could still be correct for paradigmatic emotional experiences, in particular those characteristic for basic emotions (Ekman, 1992) such as fear, anger, or disgust. The present findings suggest, however, that this objection is invalid at least for one basic emotion, disgust, for the unpleasant pictures used in the present Experiments 1a, 1b and 2 were all disgusting in nature, as was a subset of the pictures presented in Experiment 3.

Directions for Future Research

We have recently replicated and extended the findings reported in this article with another temporal judgment paradigm, the “rotating spot” method (Libet et al., 1983;

Pockett & Miller, 2007), a modern version of Wundt's (1897) and Geiger's (1903) rotating clock hand (Franikowski et al., 2021). Future studies could apply these temporal judgment methods to other emotional experiences; surprise, amusement, and curiosity are promising candidates. Furthermore, temporal judgments of affect could be combined with measurements of the physiological and facial reactions evoked by emotional stimuli (e.g., Codispoti et al., 2006; Lang et al., 1993) to study the temporal relations between subjective experience and the bodily manifestations of emotions on the individual level. Finally, the temporal judgment paradigms could be adapted to measure the subjective occurrence times of other emotion-related mental events, such as felt action tendencies. This could provide new ways of testing other prominent hypotheses of emotion psychology, such as the hypothesis that emotional experiences are (e.g., Arnold, 1960; Frijda, 1986) or cause (e.g., Mellers, 2000; Weiner, 1995) action tendencies.

Context of the Research

The ideas for the experiments reported in this article originated from our reading of a classic *Psychological Review* article by Taizo Nakashima (1909), who reported the first test of the semantic primacy hypothesis of emotion generation. While Nakashima's findings supported semantic primacy, the research had some methodological limitations. Surprisingly, although other tests of the semantic primacy hypothesis have been made, Nakashima's latency measurement approach has not been followed up until very recently. In our research, we used, for the first time, temporal order and simultaneity judgments to measure the latencies of object recognition and affect in response to affective pictures. A main advantage of these methods is that they provide relatively *pure* estimates of perceptual latencies, i.e., estimates that are uncontaminated by decisional and motor processes. Our results confirm Nakashima's conclusions. According to our findings, it takes about 120 ms from recognizing the object shown in an affective picture to experiencing the pleasant or unpleasant feeling evoked by the picture. These results support cognitive theories of emotion generation. In the meantime, we have confirmed the findings reported here with another temporal judgment method, the rotating spot method (Franikowski & Reisenzein, 2021), and with speeded reaction time measurements (Franikowski & Reisenzein, in press). The broader research program of which these studies are a part is the theoretical and empirical investigation of cognitive emotion theory, a long-standing research interest of the first author (see Reisenzein, 2019).

References

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods, 45*(4), 972–998. https://doi.org/10.3758/s13428-013-0325-2

Arnold, M. B. (1960). *Emotion and personality.* Columbia University Press.

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods, 44*(1), 158–175. https://doi.org/10.3758/s13428-011-0123-7

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379–384. https://doi.org/10.3758/BF03192707

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health, 22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Borenstein, M, Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Wiley.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276–298. https://doi.org/10.1037/1528-3542.1.3.276

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology, 39*(1/4), 106–124. https://doi.org/10.2307/1415404

Cardoso-Leite, P., & Gorea, A. (2010). On the perceptual/motor dissociation: A review of concepts, theory, experimental paradigms and data interpretations. *Seeing and Perceiving, 23*(2), 89–151. https://doi.org/10.1163/187847510x503588

Codispoti, M., Ferrari, V., & Bradley, M. M. (2006). Repetitive picture processing: Autonomic and cortical correlates. *Brain Research, 1068*(1), 213–220. https://doi.org/10.1016/j.brainres.2005.11.009

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology, 46*(4), 731–738. https://doi.org/10.1111/j.1469-8986.2009.00804.x

Colonius, H., & Diederich, A. (2011). Computing an optimal time window of audiovisual integration in focused attention tasks: illustrated by studies on effect of age and prior knowledge. *Experimental Brain Research, 212*(3), 327–337. https://doi.org/10.1007/s00221-011-2732-x

Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., Waldorp, L. J., & Wagenmakers, E.-J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review, 23*(2), 640–647. https://doi.org/10.3758/s13423-015-0913-5

Crane, T. (1992). The nonconceptual content of experience. In T. Crane (Ed.), *The contents of experience: Essays on perception* (pp. 136–157). Cambridge University Press.

Curtis, V., Aunger, R., & Rabie, T. (2004). Evidence that disgust evolved to protect from risk of disease. *Proceedings of the Royal Society London B, 271*(Suppl.), 131–133. https://doi.org/10.1098/rsbl.2003.0144

Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain.* Avon.

de Leeuw, J. R. (2014). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dimberg, U., Hansson, G., & Thunberg, M. (1998). Fear of snakes and facial reactions: A case of rapid emotional responding. *Scandinavian Journal of Psychology, 39*(2), 75–80. https://doi.org/10.1111/1467-9450.00059

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86–89. https://doi.org/10.1111/1467-9280.00221

Diederich, A. & Colonius H. (2015). The time window of multisensory integration: relating reaction times and judgments of temporal order. *Psychological Review, 122*(2), 232–241. https://doi.org/10.1037/a0038696

Dunlop, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1*(2), 170–177. https://doi.org/10.1037/1082-989X.1.2.170

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin, 53*(2), 134–140. https://doi.org/10.1037/h0045156

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe [Experimental study of the simplest mental processes]. *Archiv für die gesamte Physiologie des Menschen und der Tiere, 11*(1), 403–432. https://doi.org/10.1007/BF01659311

Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology, 2*, 243. https://doi.org/10.3389/fpsyg.2011.00243

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Lawrence Erlbaum.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116–124. https://doi.org/10.3758/bf03195503

Franikowski, P., Kriegeskorte, L.-S., & Reisenzein, R. (2021). Perceptual latencies of object recognition and affect measured with the rotating spot method: Chronometric evidence for semantic primacy. *Emotion.* Advance online publication. https://dx.doi.org/10.1037/emo0000991

Franikowski, P., & Reisenzein, R. (in press). On the latency of object recognition and affect: Evidence from speeded reaction time tasks. *Emotion.*

Frijda, N. H. (1986). *The emotions.* Cambridge University Press.

Fujisaki, W., & Nishida, S. (2009). Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research, 198*(2–3), 245–259. https://doi.org/10.1007/s00221-009-1870-x

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*(7), 773–778. https://doi.org/10.1038/nn1268

García-Pérez, M. A., & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. *Psychonomic Bulletin & Review, 19*(5), 820–846. https://doi.org/10.3758/s13423-012-0278-y

Geiger, M. (1903). Neue Complicationsversuche [New complication experiments]. *Philosophische Studien, 18*, 347–436.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7*(2), 170–184. https://doi.org/10.1016/S0959-4388(97)80005-0

Globisch, J., Hamm, A. O., Esteves, F., & Öhman, A. (1999). Fear appears fast: Temporal course of startle reflex potentiation in animal fearful subjects. *Psychophysiology, 36*(1), 66–75. https://doi.org/10.1017/s0048577299970634

Haering, C., & Kiesel, A. (2012). Mine is earlier than yours: Causal beliefs influence the perceived time of action effects. *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00393

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology, 57*(7), e13570. https://doi.org/10.1111/psyp.13570

Heath, R. A. (1984). Response time and temporal order judgement in vision. *Australian Journal of Psychology, 36*(1), 21–34. https://doi.org/10.1080/00049538408255075

Hedger, N., Adams, W. J., & Garner, M. (2015). Autonomic arousal and attentional orienting to visual threat are predicted by awareness. *Journal of Experimental Psychology: Human Perception and Performance, 41*(3), 798–806. https://doi.org/10.1037/xhp0000051

James, W. (1890). *Principles of Psychology* (Vol. 2). Dover.

Jaśkowski, P. (2014). What determines simultaneity and order perception? In V. Arstila & D. Lloyd (Eds.), *Subjective time: the philosophy, psychology, and neuroscience of temporality* (pp. 379-407). MIT Press.

Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision, 3*(7), 499–512. https://doi.org/10.1167/3.7.4

Junge, M., & Reisenzein, R. (2013). Indirect scaling methods for testing quantitative emotion theories. *Cognition and Emotion, 27*(2), 1247–1275. https://doi.org/10.1080/02699931.2013.782267

Junge, M., & Reisenzein, R. (2016). Metric scales for emotion measurement. *Psychological Test and Assessment Modeling, 58*(3), 497–530.

Kherad-Pajouh, S., & Renaud, O. (2015). A general permutation approach for analyzing repeated measures ANOVA and mixed-model designs. *Statistical Papers, 56*(4), 947–967. https://doi.org/10.1007/s00362-014-0617-3

Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: A practical introduction.* Elsevier. Retrieved from http://sub-hh.ciando.com/book/?bok_id=261990

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46*(11), 1762–1776. https://doi.org/10.1016/j.visres.2005.10.002

Knoblauch, K., & Maloney, L. T. (2012). *Modeling psychophysical data in R.* Springer. Retrieved from http://www.loc.gov/catdir/enhancemen...2943991-b.html

Kostaki, M., & Vatakis, A. (2018). Temporal order and synchrony judgments: A primer for students. In A. Vatakis, F. Balcı, M. Di Luca, & Á. Correa, (Eds.): *Timing and time perception: Procedures, measures, and applications* (pp. 233–262). Brill. https://doi.org/10.1163/9789004280205_012

Kurdi, B., Lozano, S., & Banaji, M. R. (2016). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods, 49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General, 144*(2), 339–365. https://doi.org/10.1037/xge0000040

Laird, J. D. (2007). *Feelings: The perception of self.* Oxford University Press.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist, 37*(9), 1019–1024. https://doi.org/10.1037/0003-066x.37.9.1019

Lazarus, R. S. (1995). Vexing research problems inherent in cognitive-mediational theories of emotion – and some solutions. *Psychological Inquiry, 6*(3), 183–196. https://doi.org/10.1207/s15327965pli0603_1

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaeme, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science, 1*(3), 389–402. https://doi.org/10.1177/2515245918787489

LeDoux, J. E. (1998). *The emotional brain: The mysterious underpinnings of emotional life.* Simon & Schuster.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain, 106*(3), 623–642. https://doi.org/10.1093/brain/106.3.623

Linares, D., & Holcombe, A. O. (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent. *I-Perception, 5*(6), 559–571. https://doi.org/10.1068/i0675

Linares, D., & López-Moliner, J. (2016). quickpsy: An R package to fit psychometric functions for multiple groups. *R Journal, 8*(1), 122–131. https://doi.org/10.32614/rj-2016-008

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PloS One, 8*(1), e54798. https://doi.org/10.1371/journal.pone.0054798

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* Freeman.

Meinong, A. (1894). *Psychologisch-ethische Untersuchungen zur Werttheorie. [Psychological-ethical investigations in value theory].* Leuschner & Lubensky.

Mellers, B. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin, 126*(6), 910–924. https://doi.org/10.1037/0033-2909.126.6.910

Miller, J. O., & Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: A diffusion model approach. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 394. https://doi.org/10.1037/0096-1523.32.2.394

Miller, J. O., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review, 20*(5), 819–858. https://doi.org/10.3758/s13423-013-0404-5

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393*(6684), 467–470.

https://doi.org/10.1038/30976
Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*(5), 723–739. https://doi.org/10.1037/0022-3514.64.5.723
Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from Generalized Linear Mixed-effects Models. *Methods in Ecology and Evolution, 4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x
Nakashima, T. (1909a). Time-relations of the affective processes. *Psychological Review, 16*(5), 303–339. https://doi.org/10.1037/h0075294
Nakashima, T. (1909b). Contributions to the study of the affective processes. *American Journal of Psychology, 20*(2), 157–193. https://doi.org/10.2307/1413295
Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General, 139*(2), 222–246. https://doi.org/10.1037/a0018858
Neumann, O., & Niepel, M. (2004). Timing of perception and perception of time. In C. Kaernbach, E. Schröger, & H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 245–269). Lawrence Erlbaum.
Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522. https://doi.org/10.1037//0033-295X.108.3.483
Öhman, A., & Soares, J. J. F. (1994). “Unconscious anxiety”: Phobic responses to masked stimuli. *Journal of Abnormal Psychology, 103*(2), 231–240. https://doi.org/10.1037/0021-843X.103.2.231
Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* Cambridge University Press.
Peira, N., Golkar, A., Öhman, A., Anders, S., & Wiens, S. (2012). Emotional responses in spider fear are closely related to picture awareness. *Cognition and Emotion, 26*(2), 252–260. https://doi.org/10.1080/02699931.2011.579087
Pessoa, L., Japee, S., & Ungerleider, L. G. (2005). Visual awareness and the detection of fearful faces. *Emotion, 5*(2), 243–247. https://doi.org/10.1037/1528-3542.5.2.243
Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition, 16*(2), 241–254. https://doi.org/10.1016/j.concog.2006.09.002
R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
Reisenzein, R. (2012). What is an emotion in the belief-desire theory of emotion? In F. Paglieri & C. Castelfranchi (Eds.), *The goals of cognition: Essays in honor of Cristiano Castelfranchi* (pp. 181–211). College Publications.
Reisenzein, R. (2017). Cognitive theory of emotion. In V. Zeigler-Hill & T. K. Shackleford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 1–11). Springer International.
Reisenzein, R. (2019). Cognition and emotion: A plea for theory. *Cognition and Emotion, 33*(1), 109–118. https://doi.org/10.1080/02699931.2019.1568968

Reisenzein, R., & Franikowski, P. (2022). Supplemental materials for On the Latency of Object Recognition and Affect: Evidence from Temporal Order and Simultaneity Judgments.

Reisenzein, R., & Stephan, A. (2014). More on James and the physical basis of emotion. *Emotion Review, 6*(1), 35–46. https://doi.org/10.1177/1754073913501395

Russell, J. A., & Woudzia, L. (1986). Affective judgments, common sense, and Zajonc's thesis of independence. *Motivation and Emotion, 10*, 169–183. https://doi.org/10.1007/bf00992254

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3459–3474. https://doi.org/10.1098/rstb.2009.0141

Schiller, P. H., & Kendall, J. (2004). Temporal factors in target selection with saccadic eye movements. *Experimental Brain Research, 154*(2), 154–159. https://doi.org/10.1007/s00221-003-1653-8

Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods, 17*(4), 551–566. https://doi.org/10.1037/a0029487

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology, 47*(4), 333–366. https://doi.org/10.1016/s0010-0285(03)00035-5

Siemer, M., & Reisenzein, R. (2007). The process of emotion inference. *Emotion, 7*(1), 1–20. https://doi.org/10.1037/1528-3542.7.1.1

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: When two locations are better than one. *Perception & Psychophysics, 65*(2), 318–328. https://doi.org/10.3758/bf03194803

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315. https://doi.org/10.1016/0001-6918(69)90055-9

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and perfomance IV* (pp. 629–685). Academic Press.

Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic processing precedes affect retrieval: The neurological case for cognitive primacy in visual processing. *Review of General Psychology, 10*(1), 41–55. https://doi.org/10.1037/1089-2680.10.1.41

Stumpf, C. (1899). Über den Begriff der Gemüthsbewegung [On the concept of emotion]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 21*, 47–99.

Titchener, E. B. (1905). *Experimental psychology: A manual of laboratory practice* (Vol. 2). Macmillan.

Tünnermann, J., Krüger, A., & Scharlau, I. (2017). Measuring attention and visual processing speed by model-based analysis of temporal-order judgments. *Journal of Visualized Experiments, 119*, 54856. https://doi.org/10.3791/54856 van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics, 70*(6), 955–968. https://doi.org/10.3758/PP.70.6.955

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct.* Guilford.

Weiß, K., Hilkenmeier, F., & Scharlau, I. (2013). Attention and the speed of information processing: Posterior entry for unattended stimuli instead of prior entry for attended stimuli. *PLOS ONE, 8*(1), e54257. https://doi.org/10.1371/journal.pone.0054257

Wells, F. L. (1925). Reactions to visual stimuli in affective settings. *Journal of Experimental Psychology, 8*(1), 64–76. https://doi.org/10.1037/h0066415

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*(8), 1293–1313. https://doi.org/10.3758/BF03194544

Wundt, W. (1896). *Grundriss der Psychologie [Outlines of psychology].* Engelmann.

Wundt, W. (1897). *Grundzüge der physiologischen Psychologie [Outlines of physiological psychology]* (3rd ed., Vol. 2). Engelmann.

World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. (2013). *JAMA, 310*(20), 2191. https://doi.org/10.1001/jama.2013.281053

Yarrow, K. (2018). Collecting and interpreting judgments about perceived simultaneity: A model-fitting tutorial. In A. Vatakis, F. Balcı, M. Di Luca, & Á. Correa, (Eds.), *Timing and time perception: Procedures, measures, and applications* (pp. 295–325). Brill. https://doi.org/10.1163/9789004280205_014

Yarrow, K., Jahn, N., Durant, S., & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition, 20*(4), 1518–1531. https://doi.org/10.1016/j.concog.2011.07.003

Yates, M. J., & Nicholls, M. E. (2009). Somatosensory prior entry. *Attention, Perception, & Psychophysics, 71*(4), 847–859. https://doi.org/10.3758/app.71.4.847

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175. https://doi.org/10.1037/0003-066x.35.2.151

Zheng, B., & Agresti, A. (2000). Summarizing the predictive power of a generalized linear model. *Statistics in Medicine, 19*(13), 1771–1781. https://doi.org/10.1002/1097-0258(20000715)19:13⟨1771::AID-SIM485⟩3.0.CO;2-P

Link to datasets and R scripts for the reviewers at OSF: https://osf.io/eugc5/.

Corrigendum to Publication A

Please note that this is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Aside from minor changes, we introduced some more detail and improved the writing during the copy editing process. These adaptations must not be introduced in the document due to copyright conflicts. However, major changes in comparison to this version of the document will be summarized here.

- Section **The Temporal Order of Object Recognition and Affect: Semantic or Affective Primacy?**, paragraph 1: "Cognitive [noncognitive] emotion theories [...]" was changed to: "Cognitive [noncognitive] emotion theorists [...]"
- Section **Empirical Evidence II: Reaction Time Studies**, paragraph 1: "[...], whereas the affective primacy hypothesis suggests the opposite temporal relation. In conflict with affective primacy, but supporting semantic primacy [...]" was elaborated in the final version of the manuscript: "[...]. Conversely, if the latency of object recognition is longer than the latency of affect onset, the semantic primacy hypothesis is refuted, suggesting that – in the absence of an alternative – the affective primacy hypothesis is true. Supporting semantic primacy [...]"
- Section **Objectives of the Present Research**, paragraph 3: "[...] in which the probe stimulus consists of the offset of the target stimulus (as a result, stimulus duration and probe SOA are confounded; furthermore, [...]" now reads as: "the offset of the target stimulus. Compared to the standard TOJ task, the direct reaction method has several disadvantages: The probe SOA is confounded with stimulus duration and, as a consequence, the intensity of evoked affect (which is reduced at very brief stimulus durations; see Codispoti et al., 2009; Nakashima, 1909b); furthermore, [...]"
- Section **Hypotheses and Overview**, paragraph 2: "just noticeably difference" was corrected to "just noticeable difference"
- Section **Temporal Order Judgments** of the Results section for Experiment 1a, subsection *PSSs of OR and AF*, paragraph 1: "the remaining four" was corrected to "the remaining three"
- Section **Temporal Order Judgments** of the Results section for Experiment 1a, subsection *PSSs of OR and AF*, paragraph 2: "the remaining nine" was corrected to "the remaining eight"
- Section **Meta-Analytic Integration of the Fidings and Effects of Affect Intensity**, subsection *PSS and JND*, paragraph 1: "16 ms for OR and 134 ms for OR (delay = 118 ms)" was corrected to "16 ms for OR and 134 ms for AF (delay = 118 ms)"

The final article is available, upon publication, at: https://doi.org/10.1037/xge0001244.supp

On the Latency of Object Recognition and Affect: Evidence from Temporal Order and Simultaneity Judgments

Rainer Reisenzein and Philipp Franikowski

Supplementary Materials S1: Supplementary Analyses

Analyses With All Participants Included

These analyses parallel those reported in the main article; the only difference is that participants who were excluded on theoretical (very low affect ratings) or methodological grounds (e.g., PSS beyond the SOA range; responses indicative of random or task-irrelevant responding) remain included. The sample sizes therefore correspond to those of the complete samples (N = 20 in Exp. 1a and 2; N = 36 in Exp. 3; no participants were excluded in Exp. 1b). To guard against skewed distributions and outliers, a permutation test (Kherad-Pajouh & Renaud, 2015) was computed in addition to the parametric ANOVAs.

Experiment 1a: Temporal Order Judgments

Model Fits. The average fit of the probit model to the TOJs was R^2 = .97, SD = .03 for object recognition (OR), and R^2 = .93, SD = .10 for affect (AF).

PSSs for OR and AF. The 2 (task: OR vs. AF) within-subjects × 2 (valence: pleasant vs. unpleasant pictures) × 2 (task order: OR vs. AF first) between-subjects ANOVA revealed a main effect of task, $F(1, 16)$ = 13.19, p = .002 (permutational p = .002), η_g^2 = .22. No other effect reached significance, ps ≥ .394 (permutational ps ≥ .426). The mean PSS was 40 ms for OR and 110 ms for AF; hence, the participants became on average aware of the picture-evoked feelings 70 ms after they recognized the object (compared to 79 ms for the 18 participants included in the analyses reported in the main text). One of the two excluded participants agreed with the mean pattern (PSS AF > PSS OR); the second participant (very low affect rating) had a PSS difference of -24 ms. Hence, 16 of the 20 participants had a positive OR-AF delay.

Correlations between the latencies of OR and AF. The size of the correlation between the latencies of OR and AF was r = .28 (compared to .26 for the 18 participants included in the analyses reported in the main article).

JNDs. The mean JND was greater for AF (M = 104 ms) than OR (M = 56 ms), but the difference did not reach significance, $F(1, 16)$ = 3.87, p = .467 (permutational p = .308). The other effects were also not significant, Fs ≤ 2.57, ps ≥ .769 (permutational ps ≥ .792).

Experiment 1b: Temporal Order Judgments for a broader SOA range

No participants were excluded from the analyses.

Experiment 2: Simultaneity Judgments

Model Fits. The average fit of the Gaussian function was $R^2 = .84$ for OR and $R^2 = .85$ for AF.

PSSs for OR and AF. The ANOVA revealed an effect of task, $F(1, 16) = 12.54$, $p = .003$ (permutational $p = .003$), $\eta_g^2 = .28$, reflecting that the participants became aware of the affect significantly later (219 ms) than they recognized the objects (7 ms). Thus, the participants experienced the affect 212 ms after they recognized the object (compared to 154 ms for the 16 participants included in the analyses reported in the main text). No other main effect or interaction reached significance, Fs ≤ 0.84, ps $\geq .994$ (permutational ps $\geq .997$). A positive cognition-affect lag was obtained for three of the four excluded participants; hence, 19 of 20 participants had a positive lag. The fourth excluded participant responded with "simultaneous" at all but one SOA; this participant had a negative lag of -81 ms.

Correlations between the latencies of OR and AF. The size of the Pearson correlation coefficient was reduced to $r = -.03$ (compared to .31 in the main analysis).

JNDs. The ANOVA for the JNDs (computed from the HWHH) yielded a significant effect of task type, $F(1, 16) = 9.99$, $p = .042$ (permutational $p = .007$), $\eta_g^2 = .24$, reflecting a larger spread of the psychometric function for AF (203 ms) than OR (49 ms). None of the other main or interaction effects was significant, Fs ≤ 1.64, ps $\geq .594$ (permutational ps $\geq .742$).

Experiment 3: Generalizing the Findings

Affect and Post-Experimental Ratings. Three (four) participants estimated that less than half of the pictures had induced at least a weak pleasant or unpleasant feeling in the picture preview (TOJ task). We decided to keep these participants in the analysis. The remaining participants estimated on average that 81% of the pictures induced at least a weak pleasant or unpleasant feeling (SD = 13%, range 55–95%), and that they had recognized 94% (SD = 9%, range 60–100%) of the depicted objects during the picture preview. These estimates were nearly identical to those made after the TOJ tasks, M = 96% (SD = 8%, range 55–100%) for percentage of recognized objects, and M = 82%, (SD = 15%, range 55–100%) for percentage of experienced affect.

Regarding the answer key used in case the object was not recognized, 9 participants reported no preference, 3 said they used the -1 key, 5 the +1 key, and 5 the +2 key; the remaining 12 participants recognized all objects. For trials where no affect was experienced, 7 participants reported no key preference, 4 said they used the -1 key, 10 the +1 key, and 9 the +2 key; the remaining 4 participants always experienced a feeling. For the 23 participants who reported a key preference in the AF trials, the preferred key (-1, +1, or +2) correlated weakly but not significantly with the PSS, Spearman's ρ (used because the preferred keys scale can only be considered as ordinal) = .29, $p = .177$, $n = 23$.

The participants estimated that they had felt uncertain about their response in 23% (SD = 20%, range 0–70%) of the trials in the OR task, and in 31% (SD = 21%, range 0–90%) in the AF task. A paired t-test revealed a significant difference, $t(33) = 2.82$, $p = .008$. The uncertainty ratings were essentially uncorrelated with the latencies of OR, $r(32) = .12$, $p = .511$, and AF, $r(32) = .01$, $p = .942$.

The average difficulty of the OR TOJ was $M = 2.6$ ($SD = 2.0$, range = 0–8), and that of the affect TOJ was $M = 4.8$ ($SD = 2.6$, range 0–10). The difference was significant in a paired t-test, $t(33) = 4.39$, $p < .001$. The difficulty ratings were essentially uncorrelated with the latencies of OR, $r(32) = .17$, $p = .345$, and AF, $r(32) = -.04$, $p = .820$.

Model Fits. The average fit of the probit model to the TOJs was $R^2 = .97$, $SD = .03$ for OR, and $R^2 = .92$, $SD = .12$ for AF.

PSSs for OR and AF. The 2 (task) × 2 (task order) × 2 (experiment version: online vs. lab) ANOVA of the individual PSSs revealed an effect of task, $F(1, 32) = 5.43$, $p = .026$ (permutational $p = .020$), $\eta^2_g = .06$. No other effect reached significance, Fs ≤ 2.25, ps $\geq .859$ (permutational ps $\geq .906$). The mean PSS was 56 ms for OR and 173 ms for AF; hence the participants became on average aware of the feelings 117 ms after they recognized the object (compared to 78 ms for the 34 participants included in the analyses reported in the main text). Both of the excluded participants conformed to the mean pattern (AF > OR; hence, altogether 28 of 36 conformed to the mean pattern).

Correlations between the latencies of OR and AF. The size of the correlation between the latencies of OR and AF was $r = .38$ (compared to .67 in the main article) when the outliers were included.

JNDs. The mean JND was greater for affect ($M = 140$ ms) than object recognition ($M = 65$ ms). As in the analysis reported in the main article, this difference was significant, $F(1, 32) = 13.92$, $p = .005$ (permutational $p = .007$), $\eta^2_g = .12$. None of the other effects was significant, Fs ≤ 5.17, ps $\geq .179$ (permutational ps $\geq .180$).

Analyses of Experiment 1a (TOJs) and Experiment 2 (SJs) With the Independent Channels Models for TOJ and SJ Proposed by García-Pérez and Alcalá-Quintana (GPAQ)

To check to which degree the PSS and JND estimates depended on the estimation procedure reported in the main text, we also fitted the independent channels models for TOJs and SJs proposed by García-Pérez and Alcalá-Quintana (2012; henceforth GPAQ) to the data of the first TOJ experiment (Exp. 1a) and the SJ experiment (Exp. 2). Although still infrequently used, the GPAQ model is a more realistic model of the mental processes that generate TOJs and SJs than the probit and Gaussian function models used in the main article. Therefore, but also because the GPAQ models have more parameters (up to 7), they often allow tighter fits of the psychometric curves to the data and a more precise estimate of the PSS. We refrained from interpreting the latent model parameters because our data were not precise enough for a reliable estimation of these parameters (see main text).

Brief Description of the GPAQ Models

Based on work by Sternberg and Knoll (1973) and Schneider and Bavelier (2003), García-Pérez and Alcalá-Quintana (2012; Alcalá-Quintana & García-Pérez, 2013) proposed a unified psychometric model for TOJ and SJ and provided R and Matlab routines for estimating the model's parameters. This model is a version of independent channels model discussed by Sternberg and Knoll (1973).

GPAQ-TOJ Model. In agreement with the probit model described in the main text (Exp. 1a, Method), the GPAQ model for TOJs assumes that latent judgments about temporal order are based on perceived differences in the arrival latencies of internal signals that represent the mental events of interest evoked by the target and probe stimuli. Also, in agreement with the extended version of the probit model that we fitted to the data in the main text, the GPAQ model for TOJs takes account of performance errors (e.g., pressing the wrong button) by including two response error parameters ε_1 and ε_2 that represent the probability of making an error when giving the "before" and "after" response, respectively.

There are three main differences between the probit model and the GPAQ-TOJ model. First, the GPAQ model assumes that arrival times are distributed as shifted exponential distributions. Shifted exponential distributions are often used to model arrival latencies (e.g., Colonius & Diederich, 2011; Heath, 1984), among other reasons because arrival latencies are necessarily positive, which is the case for exponentially but not for normally distributed variables. Second, the GPAQ model assumes that arrival time distributions for the target and probe perception can have different variances, which for exponentially distributed variables implies that the distribution of their differences, on which the TOJs are based, can be asymmetric. Third, the GPAQ model assumes that participants can be unable to discriminate between temporally close stimuli. This assumption is modelled by estimating a threshold parameter δ that describes the person's ability to resolve small time differences. If the difference in the arrival times are less (greater) than δ, the subject judges that one signal occurred before (after) the other; if the difference lies within $-\delta$ and $+\delta$, the subject is assumed to guess "before" with probability ξ and "after" with probability $1 - \xi$.

In all, the GPAQ model for the TOJ task contains seven parameters. Four parameters describe the process that generates the latent temporal order judgments. These are two parameters λ_1, λ_2 that describe the decay rates of the two exponential arrival time distributions (their variance thus being $1/\lambda_1$ and $1/\lambda_2$), a parameter τ that represents the difference in the arrival times of the two signals, a threshold parameter δ, and a guessing or response bias parameter ξ. In addition, there are two response error parameters ε_1 and ε_2. However, one or both of these error parameters may not be needed to model a given data set.

GPAQ-SJ Model. The GPAQ model for SJ is identical to the model for the TOJ task up to and including the computation of the decision variable (the difference in the arrival times of mental events of interest evoked by the probe and target events). The difference to the GPAQ-TOJ model results from the fact that in the SJ task, the perceived temporal differences are mapped onto different response options. In the two-alternatives SJ task used in Experiment 2, the mapping is as follows: If the differences in arrival latencies are within $-\delta$ and $+\delta$ (the lower and upper detection threshold), the subject decides that the two compared events are simultaneous, whereas if the differences in arrival latencies are greater or less than δ, the subject decides that they are non-simultaneous, and these decisions are then made public by pressing the appropriate keys. Because no guessing is necessary in the SJ task, a guessing parameter is not estimated. However, as in the TOJ task, the participant may make a response error when reporting the judgment. As it turns out, three different kinds of reporting errors are possible in the SJ task; hence three parameters for the probabilities of these different response errors ε_1, ε_2, ε_3 need to be estimated (see García-Pérez & Alcalá-Quintana, 2012). Therefore, the SJ model also has seven parameters. Again, however, one to all response error parameters may not be necessary for obtaining a good fit to a given data set.

Analyses of the Data of Experiments 1a and 2. The GPAQ model for TOJ was separately fitted to the OR and AF judgments of each participant of Experiment 1a, while the GPAQ model for SJ was fitted to the data of Experiment 2. The R package *FitTimingJudgments* (Alcalá-Quintana & García-Pérez, 2013) was used for estimation. To reduce the danger of fitting problems, we restricted the bounds of the τ and δ parameters to slightly beyond the respective SOA range (by default, these parameters are completely unrestricted). In more detail, the parameter bounds and starting values were specified in R as shown in Figure S1.6.

To reduce the danger of overfitting, models with zero, one, or two (TOJ) or zero to three (SJ) response process parameters were estimated and the model with the lowest BIC was selected (this is a recommended analysis option; see Alcalá-Quintana & García-Pérez, 2013).

Figure S1.6

Parameter Bounds and Starting Values for the GPAQ Models for Experiments 1a and 2

```
# Parameter bounds
LamBounds = c(1/600, 1/2)
# Select based on the current estimation:
    TauBounds = c(-800, 800) # for TOJ (Exp. 1a) and SJ-AF (Exp. 2)
    TauBounds = c(-250, 250) # for SJ-OR (Exp. 2)
# Select based on the current estimation:
    DeltaBounds = c(0, 800) # for TOJ (Exp. 1a) and SJ-AF (Exp. 2)
    DeltaBounds = c(0, 250) # for SJ-OR (Exp. 2)

# Starting values
LamTStart = c(1/70, 1/3)
LamRStart = c(1/70, 1/3)
TauStart = c(-70, 70)
DeltaStart = c(20, 100)
ErrStart = c(0.05)
BiasStart = c(0.5)
```

Note: R code for the parameter configuration of the GPAQ TOJ and SJ models.

On the basis of the estimated psychometric functions, the PSSs and JNDs were calculated after controlling for the effects of guessing and response errors (see Alcalá-Quintana & García-Pérez, 2013). Because the psychometric function implied by the GPAQ models can be asymmetric, the upper and the lower JND for TOJs were averaged, while the sensitivity index (Half-Width-Half-Height, HWHH) for SJs was computed as half of the difference between the left and the right 50% point of the psychometric function (half-width), and the PSS was computed as the midpoint of this interval. To make this index (HWHH) comparable to the JND index, the same conversion as in the main text was applied (i.e., HWHH was divided by 1.746; see Fujisaki & Nishida, 2009). In addition, we again computed R^2 as a measure of absolute fit of the estimated psychometric functions. As in the ANOVAs reported in

Table S1.2
Fit Indices (R^2), PSSs and JNDs for the TOJ Task in Experiment 1a, for the Probit and GPAQ Model

	R^2			PSS			JND		
Task	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Object recognition									
Probit model	.97	.03	[.88, >.99]	42	69	[-106, 119]	59	44	[2, 186]
GPAQ model	.98	.03	[.89, >.99]	58	71	[-53, 187]	65	45	[3, 202]
Affect									
Probit model	.95	.03	[.88, .99]	121	86	[-18, 246]	80	43	[8, 144]
GPAQ model	.97	.02	[.94, >.99]	117	93	[-61, 349]	117	86	[5, 290]

Note: $N = 18$.

the main article, the Holm adjustment procedure was applied to all effects for which no hypotheses had been formulated (Cramer et al., 2016).

Experiment 1a: Temporal Order Judgments

The same 18 participants as in the analyses reported in the main text were included. Table S1.2 summarizes the individual model fits of the GPAQ model for OR and AF, as well as the PSS and JND parameters computed on the bases of the estimated psychometric functions. To facilitate comparison with the results obtained with the probit model (see main text), the corresponding results are also reproduced in the table. As can be seen, the fit of the individual psychometric functions to the data was higher for the GPAQ-TOJ model (mean $R^2 = .98$ for OR and .97 for AF) than for the probit model (mean $R^2 = .97$ for OR, .95 for AF). A repeated-measures ANOVA with task (object recognition vs. affect awareness) and estimation method (probit vs. GPAQ) as within-subject factors confirmed that the GPAQ model fitted the data better than the probit model, $F(1, 17) = 13.21$, $p = .006$, $\eta^2_g = .06$. The main effect of task and the interaction were not significant, Fs ≤ 2.67, ps $\geq .241$. Thus, the model fits were consistently higher for the GPAQ model but similarly high for both tasks.

Despite the better fit of the GPAQ-TOJ model, both estimation methods yielded nearly identical estimates of the mean PSS (see Table S1.2). An ANOVA of the PSSs found that the effect of task was highly significant, $F(1, 17) = 10.93$, $p = .004$, $\eta^2_g = .16$. The main effect of estimation method and its interaction with task did not reach significance, Fs ≤ 3.08, ps $\geq .195$. Hence, for both OR and AF, the mean PSS did not depend on the estimation method (probit or GPAQ). In addition, the PSSs of the individual participants estimated by the probit model (see main text) and the GPAQ-TOJ model were highly correlated, $r = .85$ for OR and $r = .85$ for AF.

A parallel ANOVA was also conducted for the JNDs (Table S1.2). This analysis revealed a significant main effect of task, $F(1, 17) = 7.39$, $p = .044$, $\eta^2_g = .10$. The main and interaction effect of the estimation method were not significant, Fs ≤ 4.49, $p \geq .059$. Hence, the non-significant difference between the JNDs of OR and AF already found in the probit analysis (main text) was replicated in the GPAQ analysis. Hence, participants had lower JNDs for the detection of the time point of OR than

Table S1.3

Fit Indices (R^2), PSSs and JNDs for the SJ Task in Experiment 2, for the Gaussian Function and GPAQ-SJ Model

	R^2			PSS			JND		
Task	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Object recognition									
Gaussian model	.83	.18	[.38, .99]	11	24	[-14, 64]	53	22	[19, 96]
GPAQ model	.87	.14	[.57, > .99]	17	28	[-11, 81]	47	21	[14, 88]
Affect									
Gaussian model	.88	.08	[.68, .99]	165	149	[9, 512]	152	87	[45, 393]
GPAQ model	.92	.09	[.68, > .99]	159	154	[10, 540]	144	72	[29, 294]

Note: $N = 16$.

AF.

Experiment 2: Simultaneity Judgments

The average model fits, PSSs and JNDs for the 16 included participants are shown in Table 1.3. The average fits of the psychometric functions were again higher for the GPAQ model than for the probit model for both OR ($R^2 = .83$ for Gaussian function model and .87 for the GPAQ model) and AF ($R^2 = .88$ for the Gaussian function and .92 for the GPAQ model). An ANOVA with task and estimation method as within-subject factors revealed a significant main effect of estimation method, $F(1, 15) = 15.37$, $p = .004$, $\eta^2_g = .03$. The main effect of task and its interaction with estimation method were not significant, Fs ≤ 2.09, ps $\geq .338$.

A parallel ANOVA on the PSSs revealed a significant effect of task, $F(1, 15) = 15.44$, $p = .001$, $\eta^2_g = .33$, indicating that the participants became aware of the affect elicited by the pictures significantly later than they recognized the objects. No other main effect or interaction reached significance, Fs ≤ 1.47, ps $\geq .489$. The correlation between the individual PSSs estimated by the Gaussian function model used in the main text and the GPAQ-SJ model was $r = .56$ for OR and $r = .98$ for AF.

A parallel ANOVA for the JNDs yielded a significant effect of task, $F(1, 15) = 32.31$, $p < .001$, $\eta^2_g = .43$, reflecting a broader range of perceived simultaneity for AF than for OR. The effect of estimation method and the interaction were not significant, Fs ≤ 3.36, ps $\geq .174$.

Discussion

For both experiments 1a and 2, the GPAQ model had a significantly better fit than the probit or Gaussian model. The main reason was that the TOJ and SJ curves were often somewhat asymmetric (see figures S1.1 and S1.3), which could be captured by the GPAQ models but not the probit or the Gaussian function model. However, for both TOJ and SJ, the average of the PSSs for OR and AF estimated by the two models were very similar. In addition, the individual PSSs estimated with the probit/Gaussian function and the GPAQ-TOJ/SJ models correlated strongly in both experiments, with the exception of the PSS of OR in the SJ paradigm, for which the correlation was moderate.

Additional References

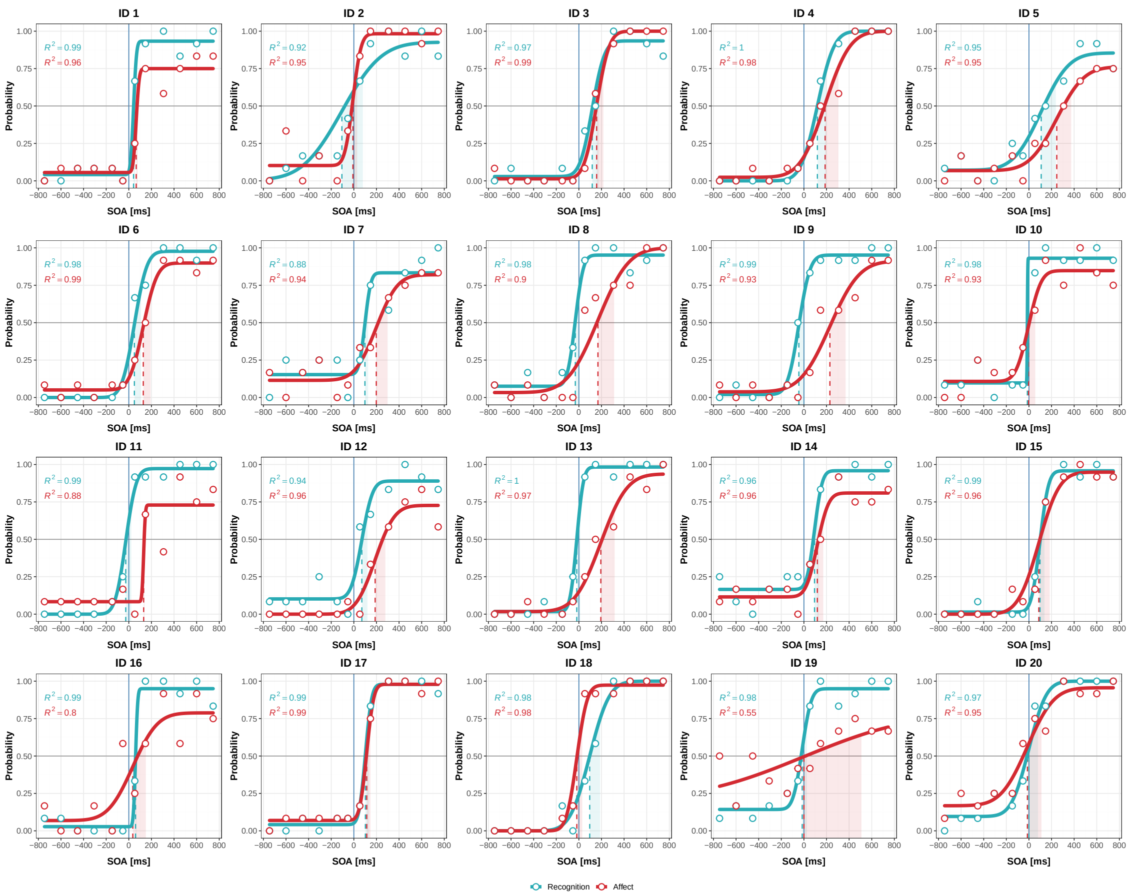
Colonius, H., & Diederich, A. (2011). Computing an optimal time window of audiovisual integration in focused attention tasks: Illustrated by studies on effect of age and prior knowledge. *Experimental Brain Research, 212*(3), 327–337. https://doi.org/10.1007/s00221-011-2732-x

Heath, R. A. (1984). Response time and temporal order judgement in vision. *Australian Journal of Psychology, 36*(1), 21–34. https://doi.org/10.1080/00049538408255075

Kherad-Pajouh, S., & Renaud, O. (2015). A general permutation approach for analyzing repeated measures ANOVA and mixed-model designs. *Statistical Papers, 56*(4), 947–967. https://doi.org/10.1007/s00362-014-0617-3

Figure S1.1

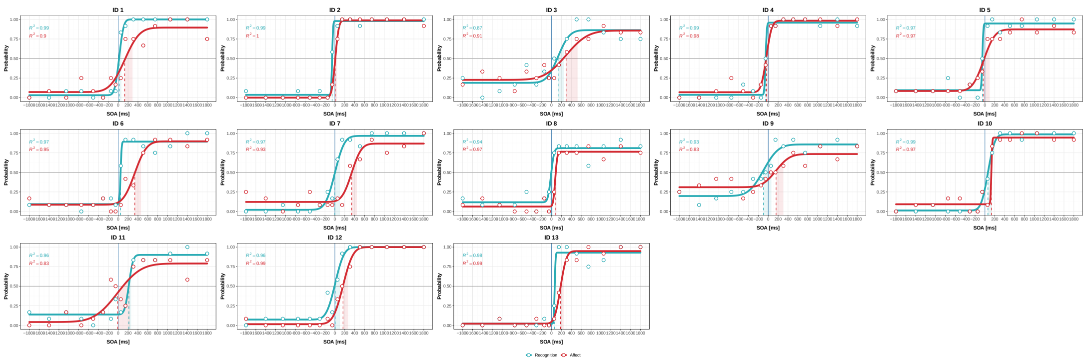
Psychometric Functions of the Individual Participants, Experiment 1a



Note: Psychometric functions for object recognition (blue) and affect (red) and the corresponding PSSs (dashed vertical line) and JNDs (shaded area under the curve) for individual participants in Experiment 1a. The fit indices (R^2) are displayed in the corresponding colors. ID 16 was excluded from the analyses reported in the main article because of very low affect ratings; ID 19 was excluded because of very high guess and lapse rate in the affect task, as a consequence of which the fitted function did not approach the asymptotes.

Figure S1.2

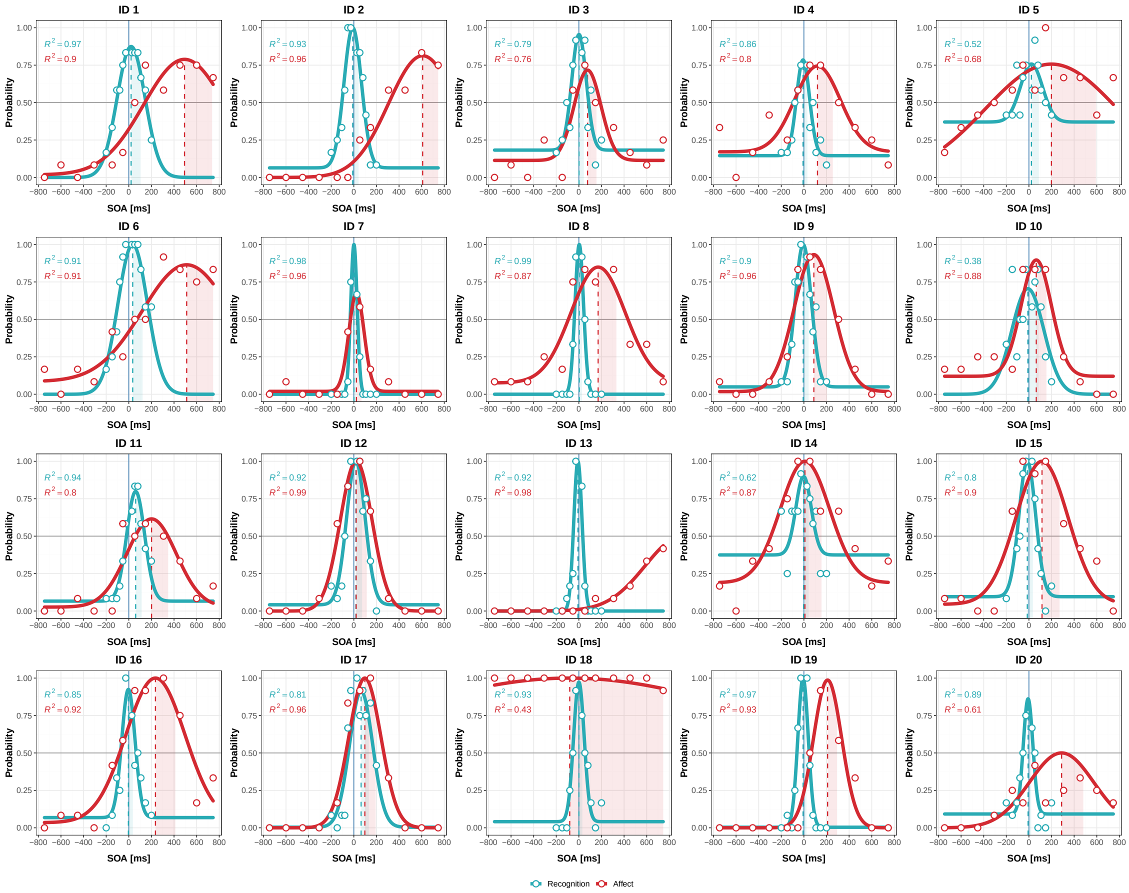
Psychometric Functions of the Individual Participants, Experiment 1b



Note: Psychometric functions for object recognition (blue) and affect (red) and the corresponding PSSs (dashed vertical line) and JNDs (shaded area under the curve) for individual participants in Experiment 1b. The fit indices (R^2) are displayed in the corresponding colors.

Figure S1.3

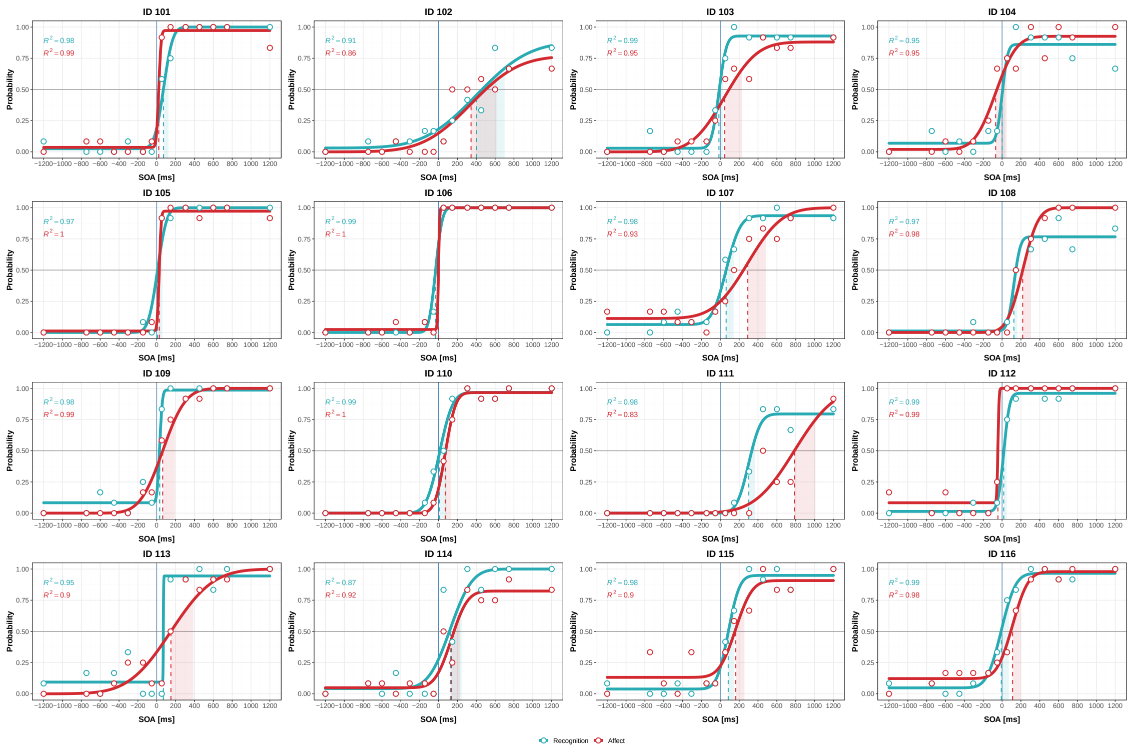
Psychometric Functions of the Individual Participants, Experiment 2



Note: Psychometric functions for object recognition (blue) and affect (red) and the corresponding PSSs (dashed vertical line) and JNDs (shaded area under the curve) for individual participants in Experiment 2. The fit indices (R^2) are displayed in the corresponding colors. ID 2 was excluded from the analyses reported in the main article because of very low affect ratings; ID 18 was excluded because affect was judged as simultaneous to the perception of the probe at all but one SOA; ID 13 was excluded because the probability of the "simultaneous" judgment failed to decline until the end of the measurement period, meaning that the estimated peak of the curve fell outside the SOA range. ID 20 was excluded because the probability of "simultaneous" judgments in the affect task remained below .50 at all SOAs, suggesting random responding. See text for further explanation.

Figure S1.4

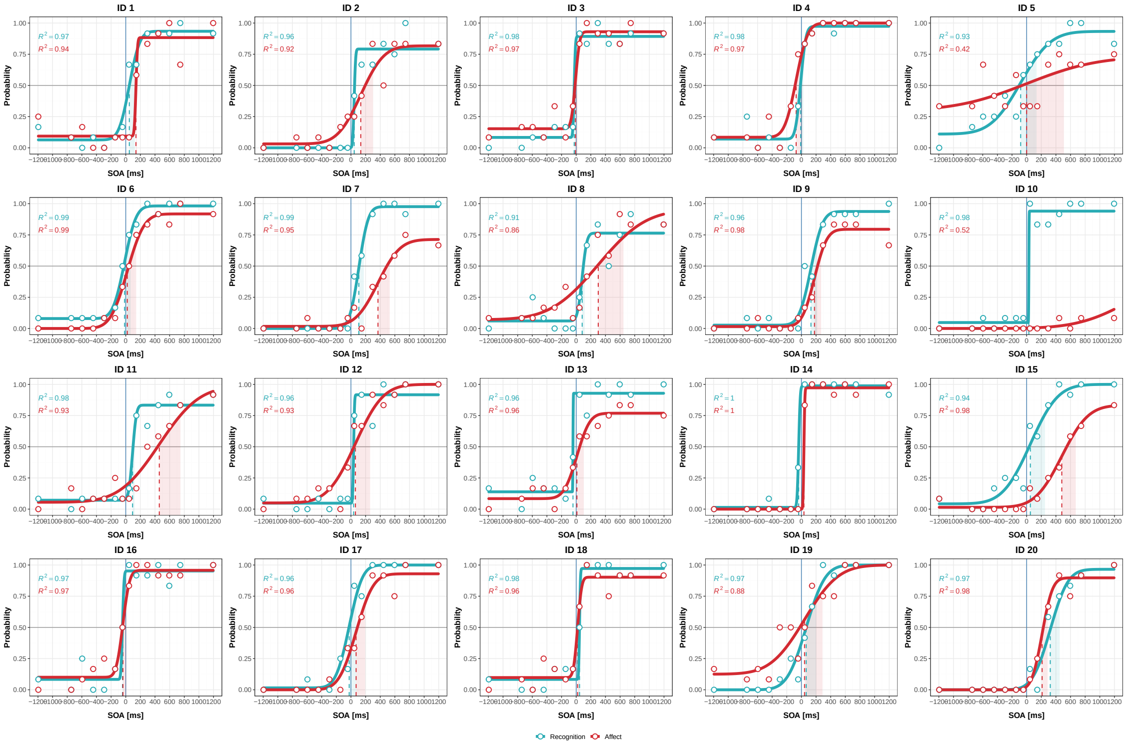
Psychometric Functions of the Individual Participants, Lab Version of Experiment 3



Note: Psychometric functions for object recognition (blue) and affect (red) and the corresponding PSSs (dashed vertical line) and JNDs (shaded area under the curve) for individual participants in Experiment 3, lab version. The fit indices (R^2) are displayed in the corresponding colors.

Figure S1.5

Psychometric Functions of the Individual Participants, Online Version of Experiment 3



Note: Psychometric functions for object recognition (blue) and affect (red) and the corresponding PSSs (dashed vertical line) and JNDs (shaded area under the curve) for individual participants in Experiment 3, online version. The fit indices (R^2) are shown in the corresponding colors. ID 5 was excluded because of a low curve fit and very high lapse and guess rates in the affect task; ID 10 was excluded because of a low curve fit and because response frequencies below 50% at all SOAs in the affect task.

Table S1.1
OASIS Pictures used in Experiment 3

Pleasant		Unpleasant	
Baby 1	Fireworks 6	Angry face 3	Gun 3
Baby 2	Flowers 4	Angry face 4	Gun 5
Baby 3	Flowers 6	Animal carcass 1	Gun 6
Baby 4	Galaxy 2	Animal carcass 3	Gun 7
Baby 5	Galaxy 5	Animal carcass 6	Gun 9
Baby 6	Galaxy 7	Baby 7	Hallway 1
Baby 9	Horse 1	Bloody knife 2	Hangover 1
Beach 1	Lake 1	Car accident 4	Injury 2
Beach 2	Lake 10	Car crash 3	Jail 1
Beach 4	Lake 12	Cemetery 5	Jail 2
Beach 6	Lake 13	Cockroach 1	Jail 3
Beach 7	Lake 14	Cockroach 2	Jail 4
Beach 8	Lake 15	Cockroach 3	Jail 5
Bird 1	Lake 2	Cockroach 4	Miserable face 2
Bird 3	Lake 3	Depressed pose 3	Miserable pose 4
Bridge 1	Lake 7	Destruction 10	Miserable pose 5
Cat 14	Lake 8	Destruction 2	Neonazi 1
Cat 3	Lake 9	Destruction 5	Pigeon 6
Cat 4	Lamb 1	Destruction 6	Plane crash 4
Cat 5	Massage 1	Destruction 7	Police 2
Children 1	Massage 2	Dog 31	Police 5
Chipmunk 2	Mother 1	Dog attack 1	Pollution 1
Dancing 5	Mother 4	Dog attack 3	Prison 2
Dancing 6	Mother 6	Explosion 2	Sad face 1
Dancing 7	Nature 1	Explosion 3	Sad face 2
Dessert 2	Nature 2	Explosion 6	Sad face 8
Dessert 3	Nude couple 5	Feces 1	Sad face 9
Dog 12	Nude couple 7	Feces 2	Shot 1
Dog 18	Penguins 1	Fence 4	Skinhead 1
Dog 19	Penguins 2	Ferret 1	Snake 1
Dog 2	Rainbow 1	Fire 3	Snake 3
Dog 20	Rainbow 2	Flood 1	Snake 4
Dog 21	School 2	Frustrated pose 1	Snake 5
Dog 3	Siblings 1	Frustrated pose 2	Soldiers 10
Dog 4	Sunflower 1	Frustrated pose 3	Soldiers 7
Dog 5	Sunset 1	Funeral 1	Soldiers 8
Dog 6	Sunset 2	Garbage dump 3	Spider 1
Dog 7	Sunset 3	Garbage dump 5	Spider 2
Elephant 1	Sunset 4	Garbage dump 8	Volcano 2
Excited face 5	Sunset 5	Gun 1	War 2
Father 1	Sunset 6	Gun 10	War 4
Fireworks 5	Zebra 1	Gun 2	War 5

Note: Pictures are sorted in an alphabetic order.

Appendix B

Publication B

Perceptual Latencies of Object Recognition and Affect Measured With the Rotating Spot Method: Chronometric Evidence for Semantic Primacy

Franikowski, P., Kriegeskorte, L.-S., & Reisenzein, R. (2022).
Emotion, 21(8), 1744–1759.
https://doi.org/10.1037/emo0000991

The final article is available, upon publication, at: https://doi.org/10.1037/emo0000991

Perceptual Latencies of Object Recognition and Affect Measured With the Rotating Spot Method: Chronometric Evidence for Semantic Primacy

Philipp Franikowski, Lara-Sophie Kriegeskorte, and Rainer Reisenzein
University of Greifswald

Abstract

According to the semantic primacy hypothesis of emotion generation, stimuli must be semantically categorized to evoke emotions. This hypothesis was tested in two chronometric studies, using the rotating spot method of timing subjective events. Participants saw pleasant and unpleasant pictures while a spot rotated around the edge of the picture. In different blocks of trials, they indicated when they experienced the pleasant or unpleasant feeling evoked by the pictures, or recognized the depicted objects, by reporting the position of the spot at the time when these mental events occurred. In both experiments, the latency of object recognition was found to be shorter than the latency of affect for nearly all participants and pictures, and the two latencies were positively correlated across participants. Experiment 2 replicated these findings and additionally showed that an experimental manipulation that delayed object recognition, blurring the pictures, also delayed the onset of affect. A mediation analysis suggested that this effect was at least partly mediated by the delayed recognition of the objects. The findings support the semantic primacy hypothesis.

Keywords: semantic vs. affective primacy; rotating spot method; mental chronometrics; latency of object recognition; latency of affective experience

Some emotional reactions seem to be evoked in an automatic, reflex-like manner by the mere perception of certain objects. For example, cat lovers involuntarily experience pleasant feelings when they see a picture of a cute cat in a magazine (Myrick, 2015), and most people are reflexively disgusted by the perception of rotten food (Junge & Reisenzein, 2013). Notwithstanding the seemingly reflex-like elicitation of these object-based feelings, cognitive emotion theorists have argued that they, like all emotions, presuppose a substantial amount of information processing including, in particular, the recognition of the object and – based on this recognition – its positive or negative evaluation (e.g., Lazarus, 1982; Ortony et al., 1988; Scherer, 2009; Storbeck et al., 2006). For example, the cat lover must recognize that the object depicted in the magazine is a cat, and must evaluate the perceived cat as being in some respect positive (e.g., as cute or aesthetically appealing). In this article, we

Author note

Philipp Franikowski, https://orcid.org/0000-0002-2109-7388, Lara Kriegeskorte, Rainer Reisenzein, https://orcid.org/0000-0002-6091-108X, Institute of Psychology, University of Greifswald
Correspondence concerning this article should be addressed to Philipp Franikowski, Institute of Psychology, University of Greifswald, Franz-Mehring-Str. 47, 17489 Greifswald. E-mail: philipp.franikowski@uni-greifswald.de
All analyzed data and R scripts are available on https://osf.io/eugc5/.

are concerned with the first of these assumptions: To evoke affect, objects must be recognized as being of a particular conceptual type; or in other words, they must be subsumed under a concept, or classified as an instance of a semantic category. This hypothesis has been termed the *semantic primacy hypothesis* of emotion generation (e.g., Storbeck et al., 2006).

Intuitively plausible as the semantic primacy hypothesis may seem to be, it has been contested by several prominent emotion theorists (e.g. LeDoux, 1998; Öhman & Mineka, 2001; Zajonc, 1980; and much earlier, Cannon, 1927; James, 1890; Watson, 1919). According to these theorists, at least some emotions, including in particular object-evoked pleasant and unpleasant feelings, can be (and perhaps even, regularly are; Zajonc, 1980) evoked by stimuli via a “noncognitive” pathway, that is, a route of information processing that bypasses higher cognitive processes including object recognition and recognition-based evaluations. This alternative theory of emotion generation is (the causal version of) the so-called *affective primacy hypothesis* (Zajonc, 1980). According to the, in our view, most plausible explication of this hypothesis, “noncognitively” caused emotions are evoked by *non- or preconceptual representations* (see e.g., Crane, 1992; Marr, 1982) of the eliciting objects, such as in the case of disgust, the perception of particular textures or color patterns characteristic of potentially infectious things; or in the case of fear, the perception of shapes characteristic of certain dangerous animals, such as “sinusoidal shapes related to snakes” (Öhman & Mineka, 2001, p. 487; see also, LeDoux, 1998). Hence, according to this theory, if you are disgusted by a moldy piece of bread, you are not disgusted because you recognized that the object in front of you is a decaying piece of bread, but because a pattern-recognition system in your brain has detected an object with a texture and color pattern characteristic of decaying organic substances.

Empirical Evidence I: Subliminal Perception Studies

The affective versus semantic primacy debate has stimulated a large number of empirical studies whose aim was to either demonstrate, or disprove, that affective stimuli can evoke emotions even if they are not recognized (including *before* they are recognized). The most-used experimental paradigm to investigate this issue is the subliminal perception paradigm. In this paradigm, affective stimuli (most often, pleasant and unpleasant pictures) are presented very briefly and then immediately backward-masked to interrupt their further processing. The assumption underlying this procedure is that, with judicious timing ($\sim$30 ms stimulus presentation time), it prevents the recognition of the depicted objects, but still allows the creation of the pre-conceptual object representations that, according to the theory of noncognitive (or better: pre-conceptual) emotion generation, evoke emotional reactions.

Seemingly supporting the affective primacy hypothesis, numerous studies found that subliminally presented affective stimuli can have effects suggestive of the elicitation of emotions: They can bias the evaluation of subsequent stimuli (affective priming: e.g., Murphy & Zajonc, 1993), and they can evoke peripheral physiological reactions (e.g., Öhman & Soares, 1994), emotional facial expressions (e.g., Dimberg et al., 2000) and activity in the amygdala (e.g., Morris et al., 1998) similar to what is found if the stimuli are presented above threshold (for a recent summary of this research, see Lähteenmäki et al., 2015).

The standard response by cognitive emotion theorists to these studies has been to accept the findings, but to emphasize that the conceptual recognition of objects, as well as their evaluation as good or bad, can occur at an unconscious level of processing, and that the findings therefore are no convincing refutation of the cognitive theory of emotion generation (e.g., Lazarus, 1995; Storbeck et al., 2006). In addition, more recent studies suggest that the stimulus presentation times used in the subliminal perception experiments were not short enough to prevent at least partial awareness of the stimuli in a subset of participants or trials (e.g., Lähteenmäki et al., 2015; Pessoa et al., 2005). Experiments that used a more sensitive psychometric threshold detection procedure and more sensitive measures of object recognition found that affective pictures elicited physiological reactions (e.g., Peira et al., 2012; see also, Codispoti et al., 2006), and could be reliably judged as pleasant or unpleasant (Lähteenmäki et al., 2015), only at presentation times at which the depicted objects were at least partly consciously recognized. These findings suggest that the occurrence of object-evoked emotions may not just require the *recognition* (semantic categorization) of the eliciting objects, but even their *conscious* recognition (for additional evidence from the continuous flash suppression paradigm, see Hedger et al., 2015).

Empirical Evidence II: Chronometric Studies

This conclusion is supported by the findings of a second, if much less abundant set of studies, in which the time needed to experience object-evoked feelings was compared to the time needed to consciously recognize the objects. The basic premise of these chronometric studies is that causes precede their effects (e.g., Mackie, 1974). Therefore, if the semantic primacy hypothesis is correct, then the latency of object recognition must be shorter (although perhaps only minimally shorter) than the latency of affect. Conversely, if the affect evoked by objects occurs before the objects have been recognized, then the semantic primacy hypothesis is refuted, implying that – in the absence of a plausible third alternative – the affective primacy hypothesis is likely correct.

The first chronometric test of the semantic versus affective primacy hypotheses was conducted more than 110 years ago by Nakashima (1909a, 1909b), who already used both of the two classic methods of timing subjective events (see Neumann & Niepel, 2004): speeded reaction times (RTs), and temporal judgments. In his RT studies, Nakashima measured how long it takes to report (by a button-press, or vocally) (a) the occurrence of pleasant and unpleasant feelings evoked by simple stimuli (colors, tones, touch) and (b) the identification of the stimuli (the recognition of the colors; the classification of the tones as low, medium, or high; the classification of the touch stimuli as hard or smooth, sharp or blunt etc.). In conflict with the affective primacy hypothesis, but consistent with semantic primacy, Nakashima found that the "cognitive" judgments took consistently less time than the corresponding judgments of affect onset. Although measured by today's methodological standards, Nakashima's studies have a number of limitations, recent RT experiments by Nummenmaa et al. (2010) confirmed their findings. These authors presented affective and neutral control pictures to their participants and asked them, in different trials, to make a semantic judgment (e.g., indicate which picture contained an animal) or an affective judgment (indicate which picture was pleasant or unpleasant).[1] Across five experiments, the semantic judgment was consistently found to take less time than the affective judgment.

Although speeded RT tasks can provide valuable insights into the temporal order of mental events (Nakashima, 1909a, 1909b; Nummenmaa et al., 2010; Posner, 1986), they do have one disadvantage: They measure not only the perceptual latencies in which researchers are usually interested, but also the time needed to decide on, and execute, the motor response. In contrast, temporal judgment tasks, the second traditional method of timing mental events (Neumann & Niepel, 2004), allow to measure perceptual latencies uncontaminated by motor response and decision times (Schneider & Bavelier, 2003).

The basic idea of all temporal judgment methods is to ask participants to relate the occurrence of a mental event of interest (e.g., the perception of a tone) to that of a contemporaneously occurring comparison event (e.g., the perception of a light). The most frequently used temporal judgment tasks are *temporal order judgments* (TOJ) and *simultaneity judgments* (SJ) (see e.g., Jaśkowski, 2014; Kostaki & Vatakis, 2018; Neumann & Niepel, 2004; Sternberg & Knoll, 1973). In these tasks, a target stimulus that reliably evokes a mental state of interest is preceded and followed, at different short intervals, by a probe stimulus, and the participants are asked to report which mental event – the target event or the perception of the probe – occurred first (TOJ), or whether the target and probe event appeared to be simultaneous or not (SJ). Data from numerous trials with varying SOAs (stimulus onset asynchronies) between target and probe are collected and summarized in proportions of "target first" judgments (TOJ) or "simultaneous" judgments (SJ). These data are then used to estimate (e.g., via fitting suitable psychometric curves to the data) the *point of subjective simultaneity* (PSS), defined as the time delay between the onset of the target and the probe stimulus at which the mental events caused by these stimuli are experienced as simultaneous. If the onset time of the target stimulus is counted as time zero, the PSS directly reflects the delay between the onset of the target mental event and the perception of the probe. Thus, TOJ and SJ tasks allow to measure the relative occurrence time of a target mental event (i.e., its occurrence time relative to the probe event).

The use of temporal judgments in affect research was again pioneered by Nakashima (1909a, 1909b), who in three additional studies used what can be construed as a variant of the TOJ method, to time the feelings evoked by colors, geometric figures and affective pictures; as well as (in two of the three studies), the recognition of the affect-evoking stimuli. Consistent with the results of his RT experiments, Nakashima again obtained support for semantic over affective primacy. However, Nakashima's temporal judgment method has a number of potentially serious shortcomings when compared to standard TOJ and SJ tasks (see Reisenzein & Franikowski, 2021). Furthermore, Nakashima did not use temporal judgments to time object recognition for affective pictures, the stimuli most often used in the subliminal perception studies and in the recent RT studies on affective vs. semantic primacy (Nummenmaa et

[1] It should be noted, however, that this judgment does not directly refer to the emotional state evoked by the stimulus, but to a property of the stimulus, its valence – presumably the capacity of the stimulus to elicit pleasant or unpleasant feelings (see Russell, 2003; Meinong, 1894). To draw inferences from the latency of the valence judgments to the latency of affect onset, one must assume that the valence judgments were based, at least in the majority of cases, on the pleasant or unpleasant feelings evoked by the stimuli (see Wells, 1925).

al. 2010). These limitations were overcome by Reisenzein and Franikowski (2021), who used standard TOJ and SJ tasks to measure the perceptual latencies of object recognition and emotional feelings (pleasure and displeasure) evoked by affective pictures. Confirming Nakashima's (1909a; 1909b) conclusions and the semantic primacy hypothesis, we found that object recognition occurred reliably earlier than the onset of affect. In addition, there was a significant positive correlation between the semantic and affective perceptual latencies (Reisenzein & Franikowski, 2021).

Objectives of the Present Research

The aim of the two experiments reported in this article was to replicate and extend the findings of Reisenzein and Franikowski (2021) with the second main variant of the temporal judgment method, the *clock method* or *rotating spot (RS) task* (e.g., Geiger, 1903; Libet et al., 1983; Pockett & Miller, 2007; Weiß, Hilkenmeier, & Scharlau, 2013). The distinguishing feature of the this temporal judgment method is that the onset of the target mental event is compared with the perceptions produced by a fast-paced stimulus stream, rather than with the perception of a single probe stimulus, as in the TOJ and SJ tasks. In the classical implementation of the clock method, the stimulus stream consists of a clock hand that revolves quickly around the center of a clock face. In modern implementations of the method, the clock is simulated on a computer monitor, and the clock hand is often (but not always; e.g., Weiß et al., 2013) replaced by a rotating spot (Pockett & Miller, 2007).[2] In experiments where the timed subjective events are induced by a stimulus, as in our studies (see also, Miller et al., 2010), a trial of the RS task typically looks as follows: (1) The clock is started and the participant is asked to watch the clock hand/RS. (2) Some time later, the target stimulus is presented and the participant is asked to notice and memorize the position of the RS at the time when the target mental event evoked by this stimulus occurred. (3) After the end of the trial, the participant reports the remembered position of the RS.

The clock method was invented by Wundt (see e.g. Wundt,1897), who initially implemented it with a fast-moving pendulum and later, a rotating clock hand (first used in a study by Geiger, 1903). The pendulum and clock paradigms were extensively used in introspective psychology to time the occurrence of subjective events, mostly simple sensations (Pockett & Miller, 2007; Spence et al., 2001). In the 1980s, the clock method was adapted by Libet et al. (1983) for use in their well-known studies on the temporal relation between the intention to make a movement, and the occurrence of the associated motor readiness potential in the EEG. Because Libet et al. (1983) used a rotating spot in place of the clock hand, this version of the method has been termed the "rotating spot method of timing subjective events" (Pockett & Miller, 2007). Following Libet et al. (1983), the RS method has been used in dozens of studies for measuring the latency of intentions (e.g., Danquah et al., 2008; Pockett & Miller, 2007; see Saigle et al., 2018 for a review). Variants of the method have also been used to time mental processes in diverse other fields of psychological research (e.g., Carlson et al., 2006; Kosovicheva & Bex, 2020; Seifried et al., 2010; Weiß & Scharlau, 2010).

The first aim of our studies, then, was to replicate the findings of Reisenzein and

[2]Yet other variants of the method use a stream of letters (e.g., Soon et al., 2008) or fast-changing colors (e.g., Kosovicheva, & Bex, 2020) instead of the clock hand.

Franikowski (2021) with the RS method. This replication attempt was motivated by the assumption that the TOJ/SJ method and the RS method each have biases, but probably not exactly the same, in which case the replication of findings obtained with one method by the other method strengthens the conclusions that can be drawn from each (cf. Linares & Holcombe, 2014; and see the General Discussion for a discussion of potential shortcomings of the RS method). We predicted that, as found in our previous studies with TOJ and SJ, the latency of affect would be longer than that of object recognition (Hypothesis 1), and that the two latencies would be positively correlated (Hypothesis 2).

The second aim of our experiments was to move beyond the previous chronometric studies of semantic vs. affective primacy by conducting an *experimental test* of the semantic primacy hypothesis within the latency measurement paradigm. To achieve this goal, Experiment 2 included an experimental manipulation (blurring the pictures) that slightly delayed object recognition. We predicted that, if the semantic primacy hypothesis is correct, this manipulation would also delay the onset of affect (Hypothesis 3), and that this effect would be mediated by the delay of object recognition (Hypothesis 4).

Ethics Statement

All experiments were designed and conducted in accordance with the Code of Ethics of the World Medical Association ("World Medical Association Declaration of Helsinki," 2013). All participants in the two experiments gave their informed consent to participate prior to the start of the experiments. They were informed that they would see pleasant and unpleasant (disgusting) pictures and that they could terminate their participation at any time without negative consequences. Ethical review and approval was not required for these studies in accordance with the German legislation and with institutional requirements. In the following description of our experiments, we report all measures, manipulations, and exclusions. All analyzed data and R scripts are available on https://osf.io/kgdnh.

Experiment 1

The aim of Experiment 1 was to validate the findings of Reisenzein and Franikowski (2021) on the perceptual latencies of object recognition and affect, obtained using TOJ and SJ judgments, with the RS method. We predicted that, as found by Reisenzein and Franikowski (2021), the measured perceptual latency of object recognition would be shorter than that of affective experience (Hypothesis 1), and that the two latencies would be positively correlated (Hypothesis 2).

Method

Participants

The participants were 19 undergraduate students, 13 females and 6 males (age M = 23.3 years, SD = 2.5). They were randomly assigned to the two between-subject conditions (valence orders; see below) of the experiment. Participants were rewarded course credit. One participant was excluded from the data analyses for reasons described in the Results, reducing the final sample size to 18.

The sample size required for detecting the predicted latency difference between object recognition and affect was estimated on the basis of the findings of Reisenzein and Franikowski (2021, Exp. 1a, 1b and 2), where mostly the same stimuli were used as in the present studies. Using TOJ and SJ tasks, we found a meta-analytically combined (three studies) effect size of $\eta_g^2 = .29$ for the mean difference between the time of object recognition and the onset of affect (120 ms), and a correlation of $r = .26$ between the two latencies. To detect an effect of this size in a within-subject design at $\alpha = .05$ with a power of .80, assuming the same correlation between the latencies, a sample size of $N = 10$ would be sufficient (calculated with G*Power; Faul et al., 2007). The actual sample size of $N = 18$ detects this effect size with a power of .987 and allows to detect an effect size of .16 with a power of .80.

Stimuli

The stimuli used in the object recognition and affect tasks were 12 pleasant and 12 unpleasant (more precisely, disgusting) pictures. The pleasant and six of the unpleasant pictures were taken from Junge and Reisenzein (2016, Experiment 1). The remaining six unpleasant pictures were replaced by higher-intensity stimuli to evoke approximately equal intensities of pleasure and displeasure. The pleasant stimuli included pictures of a laughing child, a sleeping baby, a kitten, a panda bear, a sunrise at the beach and an ice cream cup; the unpleasant (disgusting) stimuli included pictures of a finger with a displaced fingernail, vomit, maggots, a cockroach, a toilet with feces, and a moldy piece of bread. Each picture was 300 pixels wide and 360 pixels high. The pictures were collected from free sources on the internet. If necessary, pictures were cut and resized.

To familiarize the participants with the RS procedure, they first completed two other, presumably simpler tasks, in which the RS method was used to time the appearance of a stimulus and the recognition of colors. Simultaneously, these tasks provided useful additional information about the reliability and validity of the RS method. Regarding validity, we predicted that the detection of stimulus appearance would be fastest (e.g., Wood et al., 2015; see already Titchener, 1905), followed by color and shape recognition, and that both processes would be faster than the generation of affect (Nakashima, 1909a; 1909b).

The stimuli used in the color recognition task were eight small (50×50 pixels) colored squares (red, violet, blue, green, yellow, orange, black, and white). In each trial, one square was presented at the center of the picture frame. Small pictures were used to prompt the participants to look at the center of the picture, similar to what was necessary for an optimal perception of the objects in the subsequent object recognition and affect tasks (cf. Nummenmaa et al., 2010). The red square was also used in the stimulus appearance detection task.

All stimuli were presented at the center of a computer monitor with a resolution of 1024×1280 pixels and a 75 Hz refresh rate. The experiment was programmed with DMDX (Forster & Forster, 2003). Responses were collected via keyboard input as described below.

Procedure

The experiment was conducted in a computer lab with one to four subjects participating at a time. Workplaces were separated by room dividers. The experiment lasted about half an hour.

Affect Ratings. To verify that the emotional pictures had the intended emotional effect, the participants first rated the pictures for the degree of pleasure or displeasure they evoked. Half of the participants rated the pleasant and the other half, the unpleasant pictures first. Within each block, the pictures were presented in an individually randomized order. Answers were given on 11-point rating scales ranging from 0 (*not pleasant* [*unpleasant*] *at all*) to 10 (*extremely pleasant* [*unpleasant*]; the German words were *angenehm* [*unangenehm*]).

Rotating Spot Tasks. In previous studies, the clock/RS method has been implemented in a variety of ways (e.g., Danquah et al., 2008; Libet et al., 1983; Pockett & Miller, 2007; Weiß et al., 2013), using different forms (round, rectangular) and sizes of the clock face, different numbers of marked clock positions, different clock hands (traditional clock hands or rotating spots), and different rotation speeds. A systematic comparative study (Pockett & Miller, 2007) suggests, however, that the RS method is fairly robust to moderate variations of these parameters, at least for the dependent variable used in this study, the participants' perception of when they made a finger movement. We come back to the question of the accuracy of the RS method in the General Discussion.

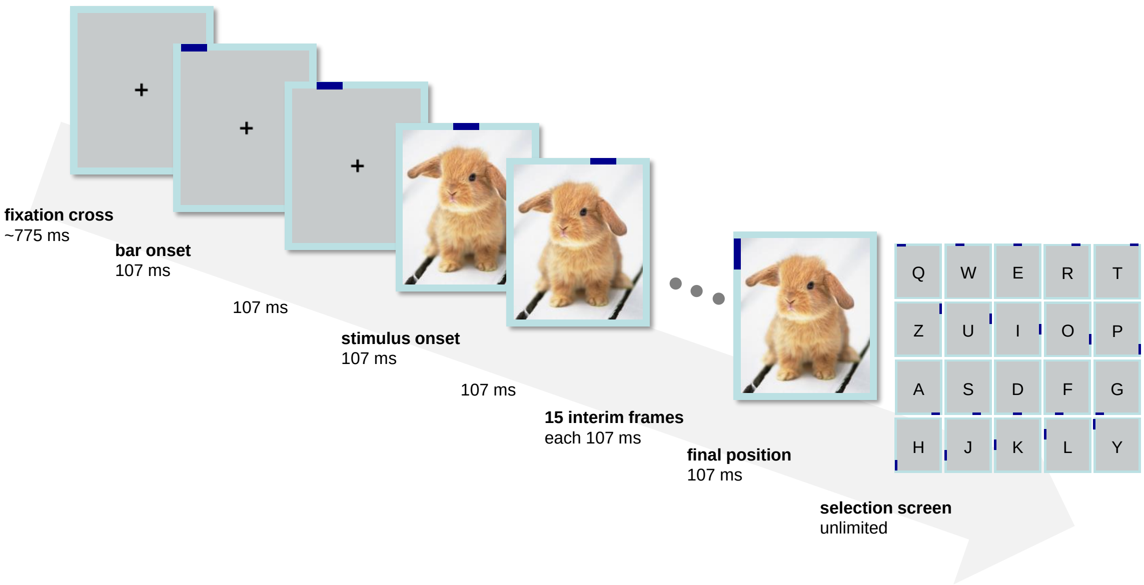
Our version of the RS method had to meet two requirements. First, it had to have sufficient temporal resolution to distinguish between the onset of object recognition and affect (about 120 ms according to Reisenzein & Franikowski, 2021, Exp. 1–2). Second, the time needed for a complete revolution of the spot (the longest interval that can be measured) had to be long enough to capture late affect onsets (about 600 ms according to Reisenzein & Franikowski, 2020) even if the pictures were presented at the latest picture onset time (+428 ms after the onset of the start of the clock; see below). We first tried a fast-moving spot (53 ms per clock tick and 20 ticks for one revolution, amounting to 1060 ms for a complete revolution), but it turned out that at this speed, the adjacent positions of our variant of the rotating spot were already difficult to distinguish. Therefore, the rotation speed was approximately halved to 107 ms per tick, which increased the measurement window to 2140 ms.[3]

To allow the participants to watch both the pictures and the RS (with peripheral vision) without gaze shifts, a rectangular clock-face was used. It consisted of a grey rectangle of the same size as the affective pictures, surrounded by a light-blue double-frame that was divided into 20 sectors, five on each side (see Figure 1). These were the possible positions of the RS. The spot consisted of a dark-blue rectangular bar that moved inside the double-frame around the edge of the picture. The "wandering bar" always started in the leftmost sector at the top of the frame and traveled clockwise around the picture until one rotation was completed (see Figure 1). Each sector of the frame was occupied by the bar for 107 ms, which amounts to 9.3 ticks per second.

[3]Although the temporal resolution of the RS was thus only 107 ms in each trial, the precision of the perceptual latency estimates is higher than that because these estimates were the average (20% trimmed mean) of the latencies obtained in numerous trials with more or less fluctuating occurrence times of the perception of the RS and the target events. These random onset fluctuations should have refined the measurement.

Figure 1

The rotating spot paradigm used in Experiments 1 and 2, illustrated with a trial from the object recognition and affect tasks



Each trial of the RS tasks looked as follows. First, a fixation cross was presented at the center of the clock face for 775 ms (Figure 1). Then, the "wandering bar" appeared on the top side of the clock and began to rotate in clockwise direction. The pictures were shown one to four clock positions (107 to 428 ms) later inside the double frame, with each picture appearing equally often at each onset position. Different picture onset times were used to avoid the development of precise expectations about the occurrence times of the pictures, which could have encouraged the use of simplifying response strategies (see the General Discussion). The participants were instructed to memorize the location of the RS when a specified mental event occurred (see below). After the spot had completed a full rotation, the picture and the clock hand disappeared and a response screen was shown. It consisted of 20 small images that showed the RS at its successive positions (see Figure 1). Each image was labeled with a letter taken from the second to fourth row of the computer keyboard, with successive letters corresponding to successive positions of the RS. The participants were asked to select the picture that showed the RS at the memorized position by entering its letter on the keyboard.

The participants completed four RS tasks in the following sequence: (1) detection of stimulus appearance, (2) recognition of color, (3) recognition of the object shown in affect-evoking pictures, and (4) onset of the pleasant or unpleasant feelings evoked by the pictures. In the stimulus appearance detection task, a red square was shown at the center of the clock face, and the participants were asked to memorize and report the position of the RS when this stimulus appeared. In the color recognition task, the participants saw one of eight differently colored squares and memorized the position of the RS when they recognized the color. In the object recognition task, they saw an affective picture and memorized the position of the RS when they recognized the depicted object. In the affect detection task, they were also shown an affective picture, but were asked to memorize the position of the RS when they experienced the pleasant or unpleasant feeling evoked by the picture. In both the object recognition and affect task, the pleasant and unpleasant pictures were presented in separate blocks whose sequence (positive or negative pictures first) was varied between participants. We presented the pleasant and unpleasant pictures in separate blocks in the attempt to simplify in particular the affect detection task, by allowing the participants to focus on the detection of a feeling of known hedonic tone.

The red square used in the appearance detection task was presented eight times at each of the four picture onset times, resulting in 32 trials for stimulus appearance. The eight colored squares were presented once at each picture onset time, resulting in 32 trials for color recognition. The 12 pleasant and the 12 unpleasant pictures were presented once at each picture onset time in both the object recognition and affect task, resulting in 96 trials (48 with positive and 48 with negative pictures) for object recognition, and 96 for affect.

Data Analysis. The reported positions of the RS when the target mental event occurred were recoded as temporal distances to the onset time of the target stimulus. That is, if the reported position of the RS corresponded to the onset time of the target stimulus, it was recoded as 0 ms; if it was one, two etc. clock positions later, it was recoded as +107 ms, +214 ms etc. If the participants performed the task as instructed, a delay of $+x$ ms thus means that the target mental event (e.g., object recognition, or affect) was experienced as occurring x ms after the perception of the spot that had been presented together with the target stimulus.

Next, the perceptual latencies of each participant were aggregated, for each of the possible combinations of tasks and the experimental control variables (see below), to a 20% trimmed mean, following a recommendation of Pockett and Miller (2007). The trimmed mean is a robust measure of location similar to the median (in fact, the median is the 50% trimmed mean); it is the mean of a variable after the bottom and top 20% of its values have been removed. A robust location measure was used because the perceptual latencies reported in the RS task typically have a right-skewed distribution similar to (although less extreme than) that characteristic for RTs, and like RTs, are usually contaminated with long-latency outliers, typically caused by lapses of attention. A good solution to both problems is the use of a robust location measure such as the median or the trimmed mean (Pockett & Miller, 2007; Wilcox, 2016). We preferred the trimmed mean because the implied averaging of several perceptual latencies yielded a more fine-grained response scale (see also Footnote 3).

The aggregated perceptual latencies were analyzed with repeated-measures analysis of variance (ANOVA) (Hypothesis 1) and correlations (Hypothesis 2). Although both hypotheses are directional, we took a conservative approach and used two-sided tests. The degrees of freedom of the ANOVA F-tests were Huynh-Feldt corrected when necessary to protect against violations of sphericity. The p-values of the effects involving experimental control variables, but not the task factor, were Holm-corrected (Holm, 1979) to protect against the inflation of the significance level in multiway ANOVA (Cramer et al., 2016). Generalized η^2 (η^2_g, Bakeman, 2005) was used as the effect size measure for the ANOVAs. All analyses were conducted in *R* (R Core Team, 2019). The add-on package *afex* was used for the ANOVAs (Singmann et al., 2020).

Results

One of the 19 participants was excluded from the data analysis because, different from all other participants, she had negative perceptual latencies for object recognition and affect paired with positive latencies of stimulus detection and color recognition, suggesting the possibility that she back-dated the perceived onset of object-recognition and affect. The cognition-affect lag of this participant was in the predicted direction. Findings changed only minimally if this participant was kept in the analysis.

Affect Ratings

The mean pleasantness ratings of the positive pictures ranged from 6.1 to 8.8 (M = 7.5, SD = 0.8) and the mean unpleasantness ratings of the negative pictures from 4.6 to 7.9 (M = 6.5, SD = 0.9). For individual participants, the mean pleasantness rating of the 12 positive pictures ranged from 6.0 to 9.0 (M = 7.5, SD = 0.9) and that of the 12 negative pictures from 2.8 to 8.6 (M = 6.5, SD = 1.6). Taken together, these findings suggest that, as intended, at least moderately intense pleasant or unpleasant feelings were induced by nearly all pictures in all participants. Reisenzein and Franikowski (2021) additionally found, for a largely identical set of pictures, that they retained their affect-evoking potential until the end of the experiment (see also Codispoti et al., 2006). Therefore, the stimuli seemed suited for timing the onset of picture-induced pleasant and unpleasant feelings with the RS method.

Latencies of Object Recognition and Affect

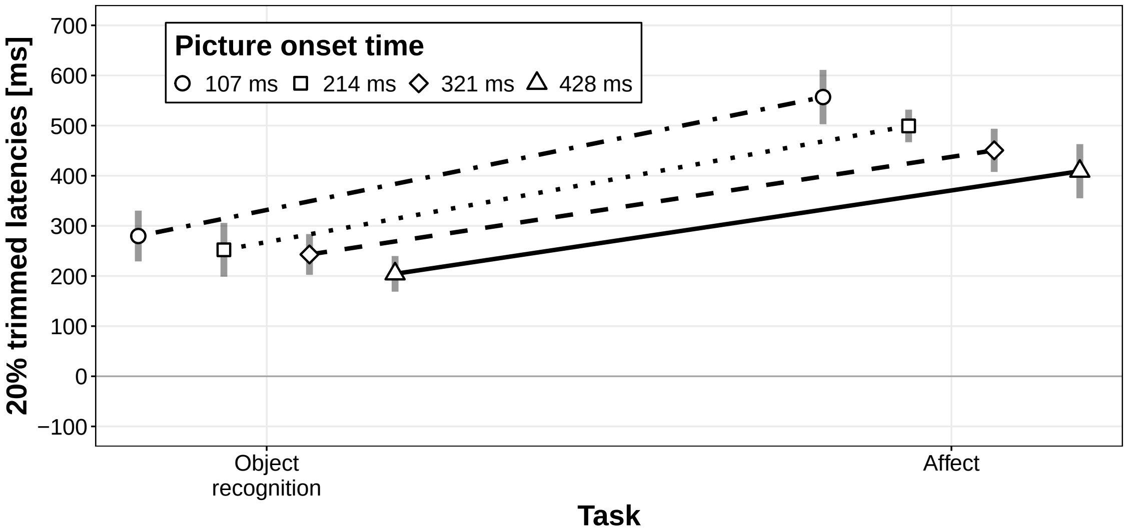
A preliminary ANOVA revealed no significant main effects or interactions for the control variables valence order and picture valence, but found significant effects for picture onset time (see Supplemental Materials S1). Therefore, picture onset time was included into the main analysis, whereas the other control variables were dropped. For the main analysis, the data were aggregated anew for each participant to 20% trimmed means in the 4 (task) × 4 (picture onset time) cells of the simplified design.

To test Hypothesis 1, we conducted a 2 (task) × 2 (picture onset time: 107, 214, 321 or 428 ms after the start of the clock) within-subjects ANOVA on the trimmed means. This analysis revealed a main effect of task, $F(1, 17) = 41.54$, $p < .001$, $\eta^2_g = .36$, as well as a main effect of picture onset time, $F(1.89, 32.2) = 38.01$, $p < .001$, $\eta^2_g = .06$, plus an interaction between task and picture onset time, $F(3, 51) = 8.29$, $p < .001$, $\eta^2_g = .01$. Because the affect latencies contained two fringe outliers (842 and 976 ms; see Figure 3), we repeated the ANOVA using a permutation test (Kherad-Pajouh & Renaud, 2015); this test replicated the results (all ps < .001).

The main effect of task confirmed Hypothesis 1, that the perceptual latency of object recognition is shorter (M = 245 ms) than that of affect (M = 470 ms; see Figures 2 and 3). However, as shown in Figure 2, this effect of task was overlayed by main and interaction effects involving picture onset time. These effects reflected the finding that the perceptual latencies became longer with earlier picture onset times, and that the effect of picture onset time was slightly stronger for affective experience than object recognition at the first two onset times (see Figure 2). The most plausible explanation of these effects is that the appearance of the RS captured the participants' attention, causing an attentional (and possibly even, a gaze) shift towards the location of the spot. Subsequently, the participants shifted attention back to the center of the clock face, where the pictures were presented; however, this return of attention was not always complete when the target stimuli occurred, which delayed their processing. This hypothesis can explain why the resulting processing delay decreased with increasing temporal distance between the start of the clock and the presentation of the target stimulus (see Figure 2). Note, however, that even at the first two picture onset times, the effects of the two factors were close to additive (see Figure 2), and that the interaction was not replicated in Experiment 2. Furthermore, follow-up simple effects analyses, conducted with one-factorial ANOVAs, confirmed that the latency of object recognition was significantly shorter than the latency of affect at each picture onset time, Fs > 31.89, ps $< .001$ (again confirmed by permutation tests; all ps $< .001$). For these reasons, the effects of picture onset time do not compromise the interpretation of the task effect.

Figure 2

Mean latencies of object recognition and affective experience at the four picture onset times, Experiment 1



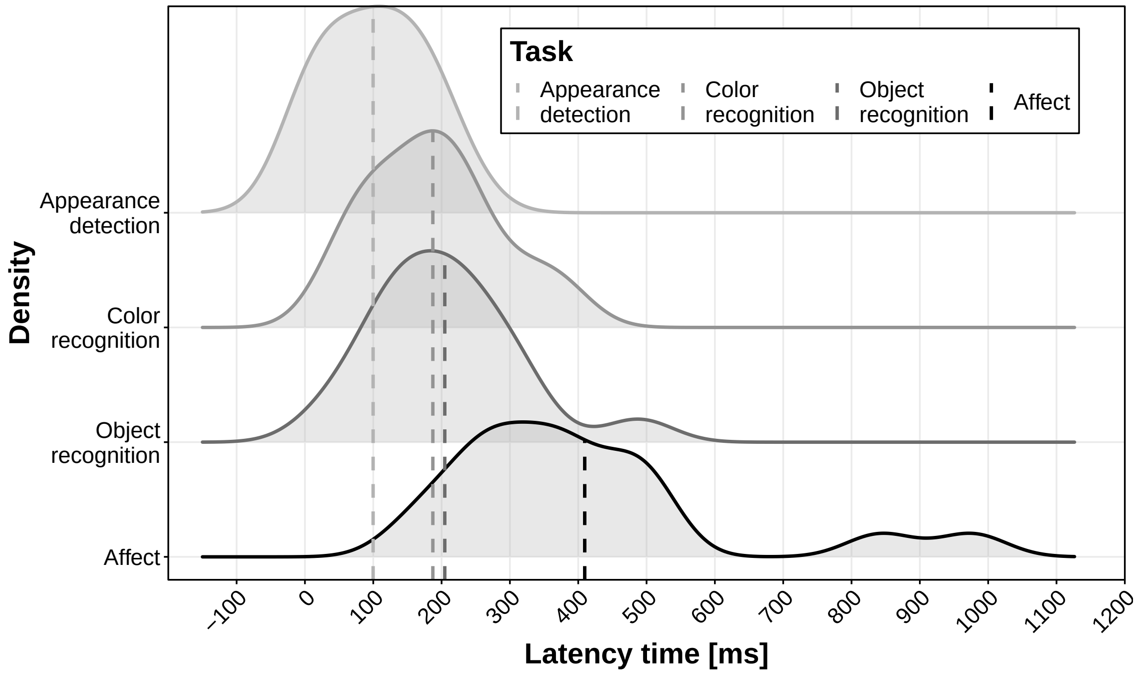
Note: Averaged 20%-trimmed means are displayed. The vertical bars show the 95%-Cousineau-Morey-CIs (Baguley, 2012) of the means.

Additional analyses for individual participants revealed that all 18 of them recognized the object shown in the affective pictures earlier than they felt the affect evoked by them.

In addition to the *temporal order* of the latencies, their size and derived from it, their *temporal distance* is of interest (Reisenzein & Franikowski, 2021). The best estimates of the size of the latencies are the measurements obtained at the latest picture onset time, where the interference caused by the start of the clock was smallest. Figure 3 shows the distributions (estimated densities) of these latencies for the four RS tasks. As expected, the average latency of appearance detection ($M =$ 101 ms) was fastest, followed by color recognition (188 ms), object recognition (204 ms) and affect (409 ms; see Figure 3). Hence, affect occurred 205 ms after object recognition. As mentioned above, the affect latencies contained two fringe outliers (Figure 3); if these are excluded, the delay reduces to 142 ms. The finding that even the simplest task (stimulus appearance detection) took 101 ms more time than the perception of the co-occurring rotating spot indicates that watching the spot still caused a delay at the latest picture onset position.

Figure 3

Estimated densities of the latencies of appearance detection, color recognition, object recognition, and affect, Experiment 1



Note: Vertical dashed lines show the means of the distributions (bandwidth = 50) of the 20%-trimmed means.

Correlation between the Latencies

Supporting Hypothesis 2, a positive and sizeable correlation between the latencies of object recognition and affect was found, $r(16) = .61$, $p = .007$. The complete correlation matrix of the four measured latencies is shown in Table 1. As can be seen, there was only one other correlation of noteworthy size, that between color and object recognition (.53); however, all correlations were insignificant after a Holm correction ($k = 5$ two-tailed tests with an overall α of .05).

To evaluate the reliability of our version of the RS method, we computed the intraclass correlation ICC(3,k) according to Shrout and Fleiss (1979; see also, Koo & Li, 2016) between the four measurements of each latency obtained at the different picture onset times. ICC(3,k) estimates the consistency of the four measurements while controlling for systematic differences between the measurement points (caused in our case by the different picture onset times), and is identical to Cronbach's α (see Shrout and Fleiss, 1979). As shown in Table 1, the reliabilities of object recognition (ICC(3, k) = .97) and affect (.99) were very high, as was the reliability of color recognition (.95), whereas the reliability of appearance detection was markedly lower (.67). A likely explanation for this finding is that the appearance detection

Table 1
Correlations between the latencies of appearance detection, color recognition, object recognition, and affect (Experiment 1)

	Appearance detection	Color recognition	Object recognition	Affect
Appearance detection	*.67*			
Color recognition	.12	*.95*		
Object recognition	.00	.53	*.97*	
Affect	.02	.24	.61**	*.99*

Note: The italicized coefficients in the diagonal are the reliabilities of the variables, computed as the intraclass correlation ICC(3,k). See text for an explanation.
** $p < .01$.

latencies varied too little (average $SD = 57$ ms) to be more reliably measured given the temporal resolution of our RS clock.

Analysis for Pictures as Units

To test whether the cognition-affect lag depended on the identity of the pictures, another 2 (task: object recognition vs. affect) × 4 (picture onset) ANOVA was conducted on the 20% trimmed means of the latencies, using pictures instead of subjects as the units of analysis (see Nummenmaa et al., 2010; Reisenzein & Franikowski, 2021). Paralleling the results of the by-subjects analysis, this ANOVA revealed main effects of task, $F(1, 23) = 888.99$, $p < .001$, $\eta^2_g = .89$ and picture onset time, $F(2.92, 67.17) = 103.19$, $p < .001$, $\eta^2_g = .52$, plus an interaction, $F(2.95, 67.74) = 20.95$, $p < .001$, $\eta^2_g = .14$. The latency of object recognition was shorter than that of affect for all 24 pictures. The average latencies (again measured at the latest picture onset time) were nearly identical to those obtained in the by-subject analysis: appearance detection 96 ms, color identification 181 ms, object recognition 198 ms, affect 383 ms. The correlation between the latencies of object detection and affect was positive as in the by-subject analysis, but nonsignificant, $r(22) = .32$, $p = .123$.

Experiment 2

Experiment 2 had two goals: To replicate the findings of Experiment 1, and to conduct an experimental test of the semantic primacy hypothesis within the latency measurement paradigm. To obtain the desired test, we looked for a manipulation that would delay but not prevent object recognition. Theoretical considerations and introspective impressions gained from comparing sharp and blurred versions of the pictures suggested that a moderate blurring of the pictures (e.g., De Cesarei & Codispoti, 2010; see also, Schupp et al., 2008) would have the desired effect. We predicted that, if the affect evoked by the pictures is indeed based on the recognition of the depicted objects, the experimental delay of object recognition will delay the onset of affect (Hypothesis 3), and the effect of blurring on affect will be statistically mediated by its delaying effect on object recognition (Hypothesis 4).

Method

Participants

The participants were 40 students, 31 females and nine males (age $M = 23.2$, $SD = 3.9$). Most were psychology undergraduates. Participants were rewarded course credit. Preliminary data analyses revealed two univariate outliers, participants with extreme affect latencies (1356 and 1103 ms; the largest latency of the remaining participants was 732 ms); these participants were excluded, resulting in a final sample size of $N = 38$.

We doubled the sample size in Experiment 2 to increase the chances of detecting the predicted effect of blurring on the latencies of object recognition and affect. A sensitivity power analysis revealed that the sample size of $N = 38$ allows to detect, with a power of .80 ($\alpha = .05$), effect sizes of η_g^2 ranging from .11 to .0055, as the correlation between the latencies of object recognition/affect for sharp and blurred pictures increases from .10 to .95. If one assumes that these latencies estimate the same latent quantities (the onset times of object recognition or affect) up to a positive lag caused by blurring, the correlation between these latencies should be as high as their reliabilities, which according to the findings of Experiment 1 were $> .95$. Based on this assumption, we could expect the F-test to detect an effect of less than 1% explained variance with a power of .80.

Stimuli

The affect-inducing pictures were mostly those used in Experiment 1, but a few unpleasant pictures were exchanged because the depicted objects (e.g. maggots, a decaying carcass) were hard to identify when the pictures were blurred. The pictures were blurred using *magick* (Ooms, 2020) by applying a Gaussian filter, a low-pass filter that results in smooth blurry image. Different grades of blurring (from σ [radius] = 7 to 11) were applied to each picture and the strongest blur grade at which the depicted object was still well recognizable was selected. Because blurring was manipulated within-subjects, the pictures were divided into two parallel sets (six pleasant and six unpleasant) matched in intensity of evoked affect, and either the first or the second set was blurred for different participants.

An additional small change from Experiment 1 was that the color recognition task was replaced by a shape recognition task. This was done to test whether the findings for color recognition generalize to shape recognition. The stimuli used in the shape recognition task were eight figures (e.g. circle, triangle, square) from the Microsoft Office™ drawing shapes library. The geometric shapes were presented as small pictures 50 × 50 pixels in size at the center of the clock face.

Design

The design of Experiment 2 comprised two factors of main interest, task (object recognition versus affect) and blurring (pictures sharp or blurred). As in Experiment 1, picture onset time, valence order (pleasant or unpleasant pictures judged first) and picture valence were varied to control for possible effects of these variables. In addition, Experiment 2 also included task order (object recognition followed by affect onset, or vice versa) and blur group (picture set A or B blurred) as control factors. Valence order, task order and blur group were varied between subjects, the remaining factors were varied within subjects. The participants were randomly assigned to the eight between-subject conditions resulting from the crossing of valence order, task order, and blur group.

Procedure and Data analysis

The procedure was largely parallel to that of Experiment 1. The participants first completed, in this order, the affect rating, the stimulus appearance detection and the shape recognition task, and then either the object recognition task followed by the affect detection task, or vice versa. The data analyses of Experiment 2 also followed the protocol of Experiment 1, making allowance for the fact that two additional hypotheses were tested. The mediation hypothesis (Hypothesis 4) was tested with multilevel path analyses, using MPlus 8.1 (Muthén & Muthén, 2017).

Results

Affect Ratings

The average affect ratings for pictures and participants were similar to those obtained in Experiment 1 (see Supplemental Materials S1). A Holm-corrected 2 (blurring: no or yes) × 2 (picture valence) within-subject × 2 (blur group) between-subject exploratory ANOVA revealed an effect of blurring, $F(1, 36) = 47.03$, $p < .001$, $\eta^2_g = .08$, as well as an effect of picture valence, $F(1, 36) = 9.04$, $p = .029$, $\eta^2_g = .07$. Blurring led to a reduction in the intensity of the affect evoked by the pictures (not blurred: $M = 7.1$, $SD = 2.5$; blurred: $M = 6.1$, $SD = 1.4$). Furthermore, pleasant pictures evoked stronger feelings ($M = 7.0$, $SD = 1.4$) than unpleasant pictures ($M = 6.1$, $SD = 1.7$). No other effect reached significance, $F \leq 5.16$, $p \geq .146$. In particular, the two matched groups of blurred pictures did not significantly differ in their capacity to induce affect.

Perceptual Latencies of Object Recognition and Affect and the Effects of Picture Blurring

A preliminary ANOVA revealed no significant main effects or interactions for any of the experimental control factors, with the exception of picture onset time (as in Experiment 1) (see Supplemental Materials S1). Therefore, picture onset time was included into the main analysis, whereas the other control variables were dropped.

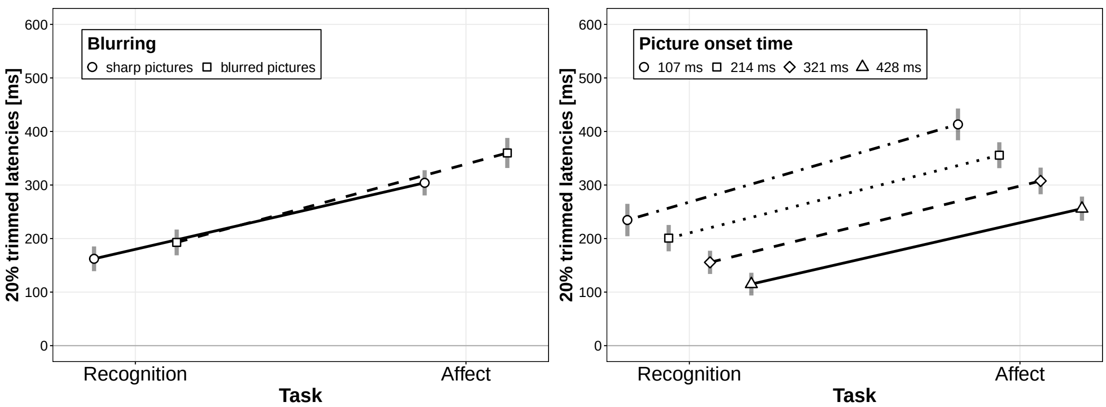
As in Experiment 1, we predicted that the perceptual latency of object recognition would be shorter than that of affect (Hypothesis 1). In addition, we predicted that blurring the pictures would delay both object recognition (manipulation check) and affect onset (Hypothesis 3). These two hypotheses were tested simultaneously in a 2 (task: object recognition or affect) × 2 (blurring: no or yes) × 4 (picture onset time) ANOVA of the 20% trimmed means, which for this purpose were newly aggregated for each participant into the 2 × 2 × 4 cells of this design. As in Experiment 1, the p-values of the predicted effects (here, the main effects of task type and blur status) were uncorrected, whereas those of the remaining main and interaction effects were Holm-corrected (Cramer et al., 2016).

The ANOVA found a main effect of task, $F(1, 37) = 31.23$, $p < .001$, $\eta^2_g = .16$, and a main effect of blurring, $F(1, 37) = 33.43$, $p < .001$, $\eta^2_g = .017$. In addition, there was a main effect of picture onset time, $F(1.98, 73.44) = 71.97$, $p < .001$, $\eta^2_g = .08$. No other effect was significant, $F \leq 3.23$, $p \geq .146$.

Supporting Hypothesis 1 and replicating Experiment 1, the latency of affect (M = 336 ms) was significantly longer than that of object recognition (M = 179 ms). Simple effects analyses using one-factorial ANOVAs confirmed that affect occurred later than object recognition at each of the four picture onset times, Fs ≥ 27.26, ps $\leq .001$, $\eta^2_g = .15$ (see Figure 4). Also as predicted, blurring the pictures delayed the recognition of the objects, M delay = 35 ms, $F(1, 37) = 27.55$, $p < .001$, $\eta^2_g = .027$; and this manipulation also delayed the onset of affect, M delay = 60 ms, $F(1, 37) = 19.38$, $p < .001$, $\eta^2_g = .019$ (Hypothesis 3; Figure 4).

Figure 4

Means of the latencies of object recognition and affective experience, Experiment 2



Note: Averaged 20%-trimmed means are displayed. The vertical bars show the 95%-Cousineau-Morey-CIs (Baguley, 2012) of the respective means. Left panel: Mean latencies for sharp and blurred pictures. Right panel: Mean latencies at the four picture onset times.

The main effect of picture onset time replicated the corresponding effect found in Experiment 1 (see Figure 4): Both object recognition and affect, were delayed by earlier picture onset times. This effect does not compromise the interpretation of the task effect, however (see also the General Discussion).

Additional analyses revealed that 31 of the 38 participants recognized the object shown in the pictures earlier than they felt the affect evoked by them (M difference = 199 ms, SD = 163 ms), whereas 7 showed the opposite pattern (M difference = 30 ms, SD = 36 ms). However, four of these latter participants also had shorter latencies for affect than for the detection of stimulus appearance. Because this finding is a priori implausible, it suggests that these participants found the affect task too difficult and therefore used a simplifying response strategy. Specifically, they may have reported the onset time of the affective picture or the time of object recognition (for further discussion, see the General Discussion).

To obtain estimates of the size of the perceptual latencies comparable to those reported in Experiment 1, we used the measurements obtained for sharp pictures at the last picture onset time. According to these estimates, the mean latency of object recognition was 95 ms, and that of affect 219 ms; hence, affect occurred 124 ms after object recognition. As in Experiment 1, the detection of stimulus appearance (92 ms) was fastest, although the difference to object recognition was only minimal.

Table 2
Correlations between the latencies of appearance detection, shape recognition, object recognition, and affect (Experiment 2)

	Appearance detection	Shape recognition	Object recognition	Affect
Appearance detection	*.80*			
Shape recognition	.27	*.89*		
Object recognition	.33	.51**	*.94*	
Affect	.32	.32	.59***	*.99*

Note: The italicized coefficients in the diagonal are the reliabilities of the variables (ICC(3,k)).
$^{**}p < .01$, $^{***}p < .001$.

Unexpectedly, the latency of shape recognition (135 ms) was even slightly greater than that of object recognition.

Correlations Between the Perceptual Latencies

Replicating Experiment 1, the predicted positive interindividual correlation between the latencies of object recognition and affect (Hypothesis 2) was again found, $r(36) = .59$, and its size was nearly identical to that found in Experiment 1 (.61). The remaining correlations among the four measured latencies were also similar to those found in Experiment 1, although some were slightly higher (see Table 2). After the Holm correction (see Exp. 1), only the correlation between object and shape recognition (.51) remained significant.

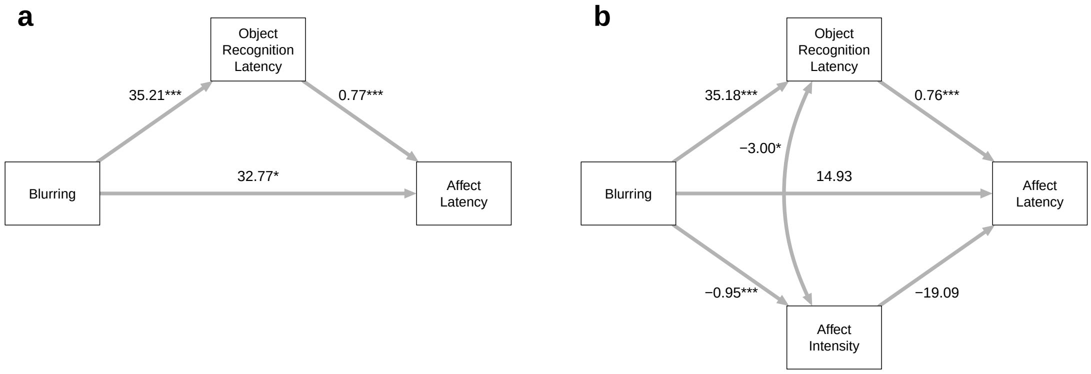
Mediation Analysis

According to Hypothesis 4, the affect evoked by blurred pictures is delayed because the depicted objects need more time to be recognized. To test this hypothesis, we conducted a mediation analysis, using the path analytic approach. The estimated path model is shown in Figure 5a.[4] Because blurring was manipulated within-subjects, the paths were estimated with multilevel structural equation modeling (Preacher et al., 2010), using Mplus 8.1 (Muthén & Muthén, 2017). To have sufficient data for estimating random slopes, the four estimates of object recognition and affect latency obtained at the four picture onset times were entered into the path analysis as replicate measurements. On the between-level (the relations between the means of the variables), the path from object recognition to affect was estimated; this path corresponds to the between-subjects correlation between object recognition and affect reported above. Although not of main interest, the inclusion of this path helped to improve model fit and thus the precision of the estimates of the paths of main interest.

[4]Although the endogenous variables of this path model are occurrence times, this is not meant to imply that the times themselves are regarded as causes or effects. Rather, the actual causes and effects are the events that occur at the respective times: Blurring delays the recognition of the object (i.e. causes it to occur at time $t + d$, rather than time t) and the delayed occurrence of object recognition, in turn, delays the occurrence of affect (causes it to occur at time $t + d + a$, rather than time $t + a$).

Figure 5

Path diagrams for the mediation analyses in Experiment 2



Note: Numbers next to the arrows are unstandardized effect sizes together with their significance levels ($^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$). Diagram **a**: the initial mediation model with object recognition latency as the mediator. Diagram **b**: the extended model including affect intensity as a second mediator.

In an attempt to maximally account for interindividual effect heterogeneity, we first tried to estimate the model shown in Figure 5a while allowing by-subject random slopes for all three paths and an unrestricted covariance structure for the random coefficients (cf. Barr et al., 2013). However, this model encountered estimation problems, indicating that it was too complex. Inspection of the variances of the random slopes obtained if only one of the three paths was allowed to vary between participants at a time, suggested that only the direct path from blurring to affect latency required a random slope. The simplified model could be estimated without problems.

The path analysis revealed significant paths from blurring to object recognition latency (unstandardized $\beta = 35$, 95% CI = [22, 48], $p < .001$ according to the critical ratio z-test), meaning that – as already found in the ANOVA – participants needed 35 ms more to recognize the object if the pictures were blurred; and from object recognition latency on affect latency ($\beta = 0.77$, 95% CI = [0.62, 0.92], $p < .001$). The indirect effect was also significant ($\beta = 27$, 95% CI = [18, 37], $p < .001$) (see Figure 5a). Unexpectedly, however, there was also a significant direct effect of blurring on affect latency ($\beta = 33$, 95% CI = [6, 59], $p = .015$).

The incomplete mediation of the blurring→affect latency effect cannot be attributed to an imperfect reliability of the mediator (Hoyle & Kenny, 1999), since the object recognition latency was highly reliable (.94; Table 2). A possible post-hoc explanation is that the effect of blurring on affect latency was in part also mediated by affect intensity: As reported above, blurring also reduced the intensity of experienced affect, and the weaker feelings may have needed slightly more time to be detected. To test this hypothesis, we estimated an extended path model that included the affect intensity ratings of the participants for sharp and blurred pictures as an additional mediator (see Figure 5b). To account for covariance between the two mediating variables not explained by blurring, the covariance between the residuals was also estimated in this model (cf. Preacher et al., 2010). To improve model fit, this parameter was additionally estimated on the between-level (the participant means), as were the paths from object recognition to affect latency, and from affect intensity to affect latency. A random slope was again specified for the direct effect of blurring on affect latency. The paths blurring→affect intensity and affect intensity→affect latency only allowed the estimation of random intercepts, because affect intensity was a constant for each participant in each experimental condition.

Comparing Figure 5a to Figure 5b reveals that the extended path model had similar-sized paths from blurring to object recognition latency (β = 35, 95% CI = [22, 48], $p < .001$) and from object recognition latency to affect latency (β = 0.76, 95% CI = [0.61, 0.91], $p < .001$), as well as a similar-sized indirect effect (β = 27, 95% CI = [17, 36], $p < .001$), whereas the direct effect of blurring on affect latency was reduced to half and no longer significant (β = 15, 95% CI = [-17, 47], $p = .364$). The hypothesized effect of blurring on affect intensity was also significant, β = -0.95, 95% CI = [-1.21, -0.69], $p < .001$. However, the effect of affect intensity on affect latency, although in the predicted direction, was not significant (β = -19, 95% CI = [-48, 10], $p = .201$), and the corresponding indirect effect of blurring on affect latency also missed significance (β = 18, 95% CI = [-11, 47], $p = .218$).

Analysis for Pictures as Units

As in Experiment 1, the latencies of object recognition and affect were also compared using pictures rather than subjects as the units of analysis. A 2 (task: object recognition vs. affect) × 4 (picture onset time) × 2 (blur) ANOVA revealed significant main effects of task, $F(1, 23) = 157.43$, $p < .001$, $\eta^2_g = .58$, picture onset time, $F(2.85, 65.47) = 221.27$, $p < .001$, $\eta^2_g = .41$, and blur status, $F(1, 23) = 14.82$, $p = .001$, $\eta^2_g = .11$. None of the interactions were significant, $F \leq 3.25$, $p \geq .110$. The latency of affect was longer than that of object recognition for 47 of the 48 (24 sharp and 24 blurred) pictures. The average latencies of object recognition (M = 170 ms) and affect (315 ms), and thus the cognition-affect delay (145 ms) were nearly identical to those obtained in the by-subjects analysis, as was the effect of blurring (+43 ms). The mean latencies of stimulus appearance detection and shape recognition were also nearly identical. As in Experiment 1, the correlation between the latencies of object recognition and affect for pictures as units was positive but not significant, $r(46) = .10$, $p = .483$.

Discussion

As in Experiment 1, object recognition was reported significantly ahead (124 ms in Exp. 2) of affective experience (Hypothesis 1), and there was a positive interindividual correlation of nearly identical size ($r = .61$) between the two latencies (Hypothesis 2). Further replicating Experiment 1, Hypothesis 1 was also supported if pictures instead of participants were used as the units of analysis, and the predicted latency differences were mostly also found on the individual level, i.e., for most participants and most pictures.

Adding to this temporal and correlational evidence for semantic primacy, Experiment 2 found that an experimental manipulation that delayed object recognition, blurring the pictures, also delayed the onset of affect (Hypothesis 3). A path analysis suggested that this effect was partly mediated by the delayed recognition of the object, providing support for Hypothesis 4. An extended path analysis provided tentative evidence that the remaining effect of blurring on affect latency may have been mediated by affect intensity.

General Discussion

The two experiments reported in this article used, for the first time, the rotating spot method of timing subjective events (Pockett & Miller, 2007) to measure the perceptual latencies of the pleasant and unpleasant feelings evoked by affective pictures, and the recognition of the objects shown in the pictures. The main goal of the experiments was to test the semantic primacy hypothesis, according to which the conceptual recognition of objects is a necessary partial cause of the affect evoked by them. Four hypotheses derived from the semantic primacy hypothesis were tested, all of which received support.

First, the perceptual latency of affect was found to be significantly longer than that of object recognition in both experiments (Hypothesis 1). On average, the delay was 205 ms in Experiment 1 and 124 ms in Experiment 2. Furthermore, the predicted delay between cognition and affect was mostly also found on the individual level, i.e., for most participants and pictures.

Second, a significant and sizeable (average $r = .60$) positive correlation between the individual perceptual latencies of object recognition and affect was obtained in both experiments in the by-subject analyses (Hypothesis 2).

Third, an experimental manipulation that delayed the recognition of the depicted objects, blurring the pictures, also increased the onset of affect (Hypothesis 3; Experiment 2).

Fourth, a mediation analysis supported the assumption that the effect of blurring on affect latency was at least in part mediated by its effect on object recognition (Hypothesis 4, Experiment 2), and there was some, although rather tentative, support for the post-hoc hypothesis that the remaining effect was mediated by affect intensity. Although the results of a statistical mediation analysis alone cannot prove a mediation hypothesis, in our case the proposed mediation model gains credibility from the fact that the causal ordering of the variables is known: Blurring was experimentally manipulated and thus had causal priority; and object recognition was shown to occur prior to affect.

Comparison to Previous Findings and Implications

The finding that the affect occurred after object recognition replicates parallel results of Reisenzein and Franikowski (2021) obtained with TOJ and SJ tasks. This finding is also in agreement with the results of previous experiments in which the latency of object recognition and affect was measured with speeded RT tasks (see the Introduction). Hence, both RT studies and temporal judgment studies using three different methods – TOJ, SJ, and the RS method – consistently support the *temporal primacy* of object recognition over affective experience. And because causes precede their effects, these findings also support the *causal primacy* of object recognition over affect.

The causal primacy of object recognition is additionally supported by the positive correlation between the latencies of object recognition and affect in the by-subjects analysis. This finding, too, replicates results by Reisenzein and Franikowski (2021) with temporal judgments (TOJ/SJ) in the by-subjects analysis, and findings of Nummenmaa et al. (2010) (Exp. 3) with RT measurements in the by-pictures analysis. The size of this correlation (on average, .60 without correction for attenuation) was actually much higher than that obtained with TOJ and SJ. This can in part be attributed to the higher reliability of the RS measurements.

Experiment 2 significantly added to the temporal and correlational evidence for semantic primacy by showing (a) that an experimental manipulation that delayed object recognition, blurring the pictures, also delayed the onset of affect; and (b) that this effect was at least partly statistically mediated by the effects of blurring on object recognition.

The conclusion that object recognition is a partial cause of the affect evoked by pictures agrees with recent findings from subliminal perception studies (Lähteenmäki et al., 2015) and the continuous flash suppression paradigm (Hedger et al., 2015). It also agrees with the findings of a few other studies that arrived at the same conclusion using yet other methods. For example, an early study by Russell and Woudzia (1986) found that the experimental facilitation of one of two possible interpretations of an ambiguous figure (the young woman/old woman figure) caused corresponding changes in the evaluation of the figure: Participants who were primed to perceive the figure as a young woman evaluated the figure as attractive, whereas those primed to perceive the figure as an old woman, evaluated it as unattractive. Presumably, these attractiveness judgments reflected the feelings evoked by the different conceptual interpretations of the figure. In another study, Schupp et al. (2008) impeded object recognition by overlaying pictures with different amounts of color pixel noise, which has a similar effect as the blurring manipulation used in our Experiment 2. Schupp et al. (2008) found that the early posterior negativity (EPN), a component of the event-related potential sensitive to emotional content, distinguished between emotional and nonemotional pictures only at color noise levels at which the objects could still be recognized. Furthermore, consistent with the findings of our Experiment 2 for the latency of affective experience, the emotional modulation of the EPN was increasingly delayed with increasing noise levels, and hence, the increasing difficulty of object recognition.

Methodological Implications

As mentioned in the introduction, the rotating spot method has been used in numerous studies to time mental events. Although the method thus seems to be well-established, its reliability and validity has been repeatedly questioned (for a recent review, see Verbaarschot et al., 2019). Our findings also have implications for this methodological debate.

A main criticism of the RS method is that, due to limitations of attentional resources (e.g., Pashler, 1994), participants cannot monitor the rotating spot and attend to the detection of the target mental event strictly at the same time, but only by switching attention between tasks, which will cause interference (e.g., Miller et al., 2011; Salter, 1989; see also, Verbaarschot et al., 2016, 2019). This dual-task interference is likely to increase random error and to introduce a systematic bias (delay) into the perceptual latencies (e.g., Salter, 1989). Under unfavorable circumstances, such as when the timed mental event has low salience, the participants may even be unable to perform the RS task as instructed. In this case, they could still try to do their best by resorting to an inferential response strategy. For example, it has been proposed that intentions to move in the Libet et al. (1983) paradigm are phenomenologically nonsalient or even unconscious, and that participants therefore try to infer their onset time from the perceived time of occurrence of the motor action (e.g., Pockett & Purdy, 2011; Dominik et al., 2017).[5]

These and related criticisms of the RS method have been investigated in several studies (e.g., Danquah et al., 2008; Dominik et al., 2017; Miller et al., 2010; Miller et al., 2011; Pockett & Miller, 2007; Sanford et al., 2017; Seifried et al., 2010; Verbarschoot et al., 2016). Taken together, the findings of these studies suggest to us that, although dual-task interference is undoubtedly present in the RS method, it does not prevent reliable and valid measurements of (relative) stimulus onset times (e.g., Seifried et al., 2010) and the perceived onset of motor movements (e.g. Pockett & Miller, 2007). In addition, there is evidence that the time needed to read a word can be reliably measured with the RS method (Kosovicheva & Bex, 2020). In contrast, the reliability and validity of RS measurement of spontaneous intentions in the Libet et al. (1983) paradigm appears to be low. Comparatively low reliability and validity has also been found for RS measurements of the time of the decision to respond in a choice RT task (Miller et al., 2010). However, it is possible that these spontaneous intentions and decisions are not just low-salience experiences, but unconscious (Pockett & Purdy, 2011). In this case, the failure of the RS method to accurately time these mental events is not a limitation of the method, but the result of its application to unfit cases.

Our results add to this body of research in two ways. First, the obtained effects of picture onset time on the perceptual latencies support the concern that watching the RS can interfere with monitoring the target mental state and cause a delay in the perceptual latencies. At the same time, however, our findings demonstrate that interference-caused delays are harmless as long as (a) interest is on relative rather than absolute perceptual latencies (e.g., on the difference between the onset time of object recognition and affect), and (b) the delays affect the compared latencies to the same degree (Exp. 2) or differential delays are too small to explain the differences between the compared latencies (Exp. 1).

Second, our findings extend the existing empirical support for the RS method by suggesting that reliable and valid measurements of perceptual latencies can also be obtained for several mental states not previously considered in RS research: color and shape recognition, the recognition of complex natural objects, and affect. The reliability of these RS latency measurements is documented by the high ICCs obtained for them. Their validity cannot be demonstrated in a similarly direct way; nevertheless, two lines of evidence provide indirect support for it. To save space, we focus on the validity of the measures of the two latencies of central interest in our studies, those of object recognition and affect.

The assumption that the reported object recognition and affect onset times are valid (up to a systematic, but innocuous delay) means that they were based on the introspection of the onsets of the respective mental events, and thus measured what we intended them to measure. The first line of evidence for this assumption is that most versions of the only proposed alternative account of how the RS reports

[5]These methodological concerns apply also to TOJ and SJ (e.g. Reisenzein & Franikowski, 2021), but probably to a lesser degree because these judgments are simpler to make (see also Verbaarschot et al., 2016, 2019).

were generated, is incompatible with our data. This alternative account is that the reported onset times of object recognition and/or affect were inferred from the onset of other, more easily detectable mental states (Pockett & Purdy, 2011; Seifried et al., 2010; note that this hypothesis must assume that at least *some* mental states – those that serve as the basis of inference – can be introspectively discerned even under conditions of divided attention). We proposed a specific version of this hypothesis ourselves to explain why the predicted differences between object recognition and affect were not found for some participants in Experiment 2: These participants may have simplified the affect detection task by (strategy 1) reporting the onset of the target picture, which presumably could be easily detected even with peripheral attention. The object recognition task could have been simplified the same way. Alternatively (strategy 2), the participants could have inferred the onset of affect from that of object recognition or vice versa.

However, none of these response strategies could have been used by all or even just the majority of the participants; for in that case, no systematic mean differences would have been found between the reported latencies of affect and/or object recognition, and the latency of stimulus appearance (which can be regarded as a good estimate for the unmeasured latency of detecting the onset of the affective pictures). It could be argued, however, that many or all participants used a more sophisticated variant of the described strategies: They may have (strategy 3) inferred the time of object recognition or affect by adding a small number of RS positions to the position of the RS at picture onset, based on the assumption that object recognition and affect must have occurred briefly after picture onset (see Seifried et al., 2010, for an analogous case); or (strategy 4) they may have infered the onset of affect from that of object recognition or vice versa, by adding or subtracting RS positions. Widespread use of the strategy 3 is, however, incompatible with the very low correlations of the object recognition and affect latencies to the stimulus appearance latencies. Only strategy 4 would be compatible with our data.

The second line of evidence supporting the validity of the RS object recognition and affect latency reports is the already mentioned, cross-method consistency of our main finding of a cognition-affect delay. This delay has also been found with TOJ and SJ (Reisenzein & Franikowski, 2021) and with speeded RT tasks (Nakashima, 1909a; 1909b; Nummenmaa et al., 2010). TOJ and SJ tasks probably involve less interference than the RS task (see also Verbarschoot et al., 2016; 2019) and should therefore cause less pressure to resort to simplifying response strategies. More compellingly, the RT studies of Nummenmaa et al. (2010) provided for an objective performance criterion, by asking the participants to make discriminatory semantic and affective judgments of pictures (contains an animal; is [un-]pleasant; Nummenmaa et al., 2010). The low error rate of these judgments indicates that at least in this RT task, the participants performed as instructed; i.e., their responses were indeed based on object recognition and affective valence. In our view, the most plausible, or at any rate the most parsimonious explanation of the finding that the same cognition-affect delay was found with the RS method is that these measurements were indeed what they were intended to be: introspective reports about the time of object recognition and affective experience.

References

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods, 44*(1), 158–175. https://doi.org/10.3758/s13428-011-0123-7

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379–384. https://doi.org/10.3758/BF03192707

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Cannon, W. B. (1927). The James–Lange theory of emotions: A critical examination and an alternative theory. *American Journal of Psychology, 39*(1), 106–124.

Carlson, T. A., Hogendoorn, H., & Verstraten, F. A. J. (2006). The speed of visual attention: What time is it? *Journal of Vision, 6*(12), 1406–1411. https://doi.org/10.1167/6.12.6

Codispoti, M., Ferrari, V., Junghöfer, M., & Schupp, H. T. (2006). The categorization of natural scenes: Brain attention networks revealed by dense sensor ERPs. *NeuroImage, 32*(2), 583–591. https://doi.org/10.1016/j.neuroimage.2006.04.180

Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., Waldorp, L. J., & Wagenmakers, E.-J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review, 23*(2), 640–647. https://doi.org/10.3758/s13423-015-0913-5

Crane, T. (1992). The nonconceptual content of experience. In T. Crane (Ed.), *The contents of experience: Essays on perception* (pp. 136–157). Cambridge University Press.

Danquah, A. N., Farrell, M. J., & O'Boyle, D. J. (2008). Biases in the subjective timing of perceptual events: Libet et al. (1983) revisited. *Consciousness and Cognition, 17*(3), 616–627. https://doi.org/10.1016/j.concog.2007.09.005

De Cesarei, A., & Codispoti, M. (2010). Effects of picture size reduction and blurring on emotional engagement. *PloS One, 5*(10), e13399.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86–89.

Dominik, T., Dostál, D., Zielina, M., Šmahaj, J., Sedláčková, Z., & Procházka, R. (2017). Libet's experiment: Questioning the validity of measuring the urge to move. *Consciousness and Cognition, 49*, 255–263. https://doi.org/10.1016/j.concog.2017.01.017

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116–124.

Franikowski, P., Kriegeskorte, L.-S., & Reisenzein, R. (2021). *Perceptual latencies of object recognition and affect measured with the rotating spot method: Further chronometric evidence for semantic primacy.* https://osf.io/kgdnh

Geiger, M. (1903). Neue Complicationsversuche. *Philosophische Studien, 18*, 347–436.

Hedger, N., Adams, W. J., & Garner, M. (2015). Fearful faces have a sensory advantage in the competition for awareness. *Journal of Experimental Psychology: Human Perception and Performance, 41*(6), 1748–1757. https://doi.org/10.1037/xhp0000127

Hoyle, R. H., & Kenny D, A. (1999). Sample size, reliability, and tests of statistical mediation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 195–222). Sage.

James, W. (1890). *The Principles of Psychology.* Holt.

Jaśkowski, P. (2014). What determines simultaneity and order perception? In V. Artstila & D. Lloyd (Eds.), *Subjective time: The philosophy, psychology, and neuroscience of temporality* (pp. 379–407). MIT Press.

Junge, M., & Reisenzein, R. (2013). Indirect scaling methods for testing quantitative emotion theories. *Cognition & Emotion, 27*(7), 1247–1275. https://doi.org/10.1080/02699931.2013.782267

Kherad-Pajouh, S., & Renaud, O. (2015). A general permutation approach for analyzing repeated measures ANOVA and mixed-model designs. *Statistical Papers, 56*(4), 947–967. https://doi.org/10.1007/s00362-014-0617-3

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kosovicheva, A., & Bex, P. J. (2020). What color was it? A psychophysical paradigm for tracking subjective progress in continuous tasks. *Perception, 49*(1), 21–38. https://doi.org/10.1177/0301006619886247

Kostaki, M., & Vatakis, A. (2018). Temporal order and synchrony judgments: A primer for students. In A. Vatakis, F. Balci, M. Di Luca, & Á. Correa (Eds.), *Timing and time perception: Procedures, measures, and applications* (pp. 233–262). Brill.

Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General, 144*(2), 339–365. https://doi.org/10.1037/xge0000040

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention, 1*, 39–58.

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist, 37*(9), 1019–1024. https://doi.org/10.1037/0003-066X.37.9.1019

Lazarus, R. S. (1995). Vexing research problems inherent in cognitive-mediational theories of emotion – and some solutions. *Psychological Inquiry, 6*, 183–196.

LeDoux, J. E. (1998). *The emotional brain: The mysterious underpinning of emotional life.* Simon & Schuster.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain, 106*(3), 623–642. https://doi.org/10.1093/brain/106.3.623

Linares, D., & Holcombe, A. O. (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent.

I-Perception, 5, 559–571. https://doi.org/10.1068/i0675

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*(3), 203–229. https://doi.org/10.1037/a0012869

Mackie, J. L. (1974). *The cement of the universe: A study of causation.* Clarendon Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* W. H. Freeman and Company.

Meinong, A. (1894). *Psychologisch-ethische Untersuchungen zur Werttheorie. [Psychological-ethical investigations in value theory].* Leuschner & Lubensky.

Meuwese, J. D. I., van Loon, A. M., Lamme, V. A. F., & Fahrenfort, J. J. (2014). The subjective experience of object recognition: Comparing metacognition for object detection and object categorization. *Attention, Perception, & Psychophysics, 76*(4), 1057–1068. https://doi.org/10.3758/s13414-014-0643-1

Miller, J., Shepherdson, P., & Trevena, J. (2011). Effects of clock monitoring on electroencephalographic activity. *Psychological Science, 22*(1), 103–109. https://doi.org/10.1177/0956797610391100

Miller, J., Vieweg, P., Kruize, N., & McLea, B. (2010). Subjective reports of stimulus, response, and decision times in speeded tasks: How accurate are decision time reports? *Consciousness and Cognition, 19*(4), 1013–1036. https://doi.org/10.1016/j.concog.2010.06.001

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393*(6684), 467–470. https://doi.org/10.1038/30976

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723–739.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Myrick, J. G. (2015). Emotion regulation, procrastination, and watching cat videos online: Who watches Internet cats, why, and to what effect? *Computers in Human Behavior, 52*, 168–176. https://doi.org/10.1016/j.chb.2015.06.001

Nakashima, T. (1909a). Time-relations of the affective processes. *Psychological Review, 16*(5), 303–339. https://doi.org/10.1037/h0075294

Nakashima, T. (1909b). Contributions to the study of the affective processes. *American Journal of Psychology, 20*(2), 157–193.

Neumann, O., & Niepel, M. (2004). Timing of perception and perception of time. In C. Kaernbach, E. Schröger, & E. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 245–269). Lawrence Erlbaum.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General, 139*(2), 222–246. https://doi.org/10.1037/a0018858

Ooms, J. (2020). *Magick: Advanced Graphics and Image-Processing in R* (Version 2.3) [Computer software].

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* Cambridge University Press.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Towards an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522.

Öhman, A., & Soares, J. J. F. (1994). "Unconscious anxiety": Phobic responses to masked stimuli. *Journal of Abnormal Psychology, 103*(2), 231–240.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Peira, N., Golkar, A., Öhman, A., Anders, S., & Wiens, S. (2012). Emotional responses in spider fear are closely related to picture awareness. *Cognition & Emotion, 26*(2), 252–260.

Pessoa, L., Japee, S., & Ungerleider, L. G. (2005). Visual awareness and the detection of fearful faces. *Emotion, 5*(2), 243–247.

Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition, 16*, 241–254. https://doi.org/10.1016/j.concog.2006.09.002

Pockett, S., & Purdy, S. (2011). Are voluntary movements initiated preconsciously? The relationships between readiness potentials, urges and decisions. In W. Sinnott-Armstrong, & L. Nadel (Ed.), *Conscious will and responsibility: A tribute to Benjamin Libet* (pp. 34–46). Oxford University Press.

Posner, M. I. (1986). *Chronometric explorations of mind.* Oxford University Press.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. https://doi.org/10.1037/a0020141

R Core Team. (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Reisenzein, R., & Franikowski, P. (2021). *On the latency of object recognition and affect: Evidence from temporal order and simultaneity judgments.* Manuscript prepared for resubmission.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172.

Russell, J. A., & Woudzia, L. (1986). Affective judgments, common sense, and Zajonc's thesis of independence. *Motivation and Emotion, 10*(2), 169–183.

Saigle, V., Dubljevi, V., & Racine, E. (2018). The impact of a landmark neuroscience study on free will: A Qualitative analysis of articles using Libet and colleagues' methods. *American Journal of Bioethics Neuroscience, 9*(1), 29–41. https://doi.org/10.1080/21507740.2018.1425756

Salter, D. (1989). Voluntary process and the readiness potential: Asking the right questions. *Behavioral and Brain Sciences, 12*(1), 181–182. https://doi.org/10.1017/S0140525X00024894

Sanford, P., Lawson, A. L., King, A. N., & Major, M. (2020). Libet's intention reports are invalid: A replication of Dominik et al. (2017). *Consciousness and Cognition, 77*, e102836. https://doi.org/10.1016/j.concog.2019.102836

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3459–3474. https://doi.org/10.1098/rstb.2009.0141

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive*

Psychology, 47(4), 333–366. https://doi.org/10.1016/S0010-0285(03)00035-5

Schupp, H. T., Stockburger, J., Schmälzle, R., Bublatzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: Brain potentials and stimulus identification. *NeuroReport, 19*(2), 167–171. https://doi.org/10.1097/WNR.0b013e3282f4aa42

Seifried, T., Ulrich, R., Bausenhart, K. M., Rolke, B., & Osman, A. (2010). Temporal preparation decreases perceptual latency: Evidence from a clock paradigm. *Quarterly Journal of Experimental Psychology, 63*(12), 2432–2451. https://doi.org/10.1080/17470218.2010.485354

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of Factorial Experiments.* R package version 0.27-2. https://CRAN.R-project.org/package=afex [Computer software].

Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience, 11*(5), 543–545. https://doi.org/10.1038/nn.2112

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130*(4), 799–832. https://doi.org/10.1037/0096-3445.130.4.799

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and perfomance IV* (pp. 629–685). Academic Press.

Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic processing precedes affect retrieval: The neurological case for cognitive primacy in visual processing. *Review of General Psychology, 10*(1), 41–55. https://doi.org/10.1037/1089-2680.10.1.41

Titchener, E. B. (1905). *Experimental psychology: A manual of laboratory practice* (Vol. 2). Macmillan.

Verbaarschot, C., Haselager, P., & Farquhar, J. (2016). Detecting traces of consciousness in the process of intending to act. *Experimental Brain Research, 234*(7), 1945–1956. https://doi.org/10.1007/s00221-016-4600-1

Verbaarschot, C., Haselager, P., & Farquhar, J. (2019). Probing for intentions: Why clocks do not provide the only measurement of time. *Frontiers in Human Neuroscience, 13*, 68. https://doi.org/10.3389/fnhum.2019.00068

Watson, J. B. (1919). *Psychology from the standpoint of a behaviorist.* Lippincott.

Weiß, K., Hilkenmeier, F., & Scharlau, I. (2013). Attention and the speed of information processing: Posterior entry for unattended stimuli instead of prior entry for attended stimuli. *PLoS ONE, 8*(1), e54257. https://doi.org/10.1371/journal.pone.0054257

Wells, F. L. (1925). Reactions to visual stimuli in affective settings. *Journal of Experimental Psychology, 8*(1), 64–76.

Wilcox, R. R. (2016). *Introduction to robust estimation and hypothesis testing* (4th ed.). Elsevier.

Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Factors influencing the latency of simple reaction time. *Frontiers in Human Neuroscience, 9.* https://doi.org/10.3389/fnhum.2015.00131

World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. (2013). *JAMA, 310*(20), 2191. https://doi.org/10.1001/jama.2013.281053

Wundt, W. (1896). *Grundriss der Psychologie [Outlines of psychology]*. Engelmann.

Wundt, W. (1897). *Physiologische Psychologie [Physiological psychology]* (3rd ed.). Engelmann.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Corrigendum to Publication B

Please note that this is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Aside from minor changes, we introduced some more detail and improved the writing during the copy editing process. These adaptations must not be introduced in the document due to copyright conflicts. However, major changes in comparison to this version of the document will be summarized here.

- Section **Empirical Evidence II: Chronometric Studies**, paragraph 5: "The use of temporal judgments in affect research was again pioneered by Nakashima (1909a, 1909b), who in three additional studies used what can be construed as a variant of the TOJ method, to time the feelings evoked by colors, geometric figures and affective pictures; as well as (in two of the three studies), the recognition of the affect-evoking stimuli. Consistent with the results of his RT experiments, Nakashima again obtained support for semantic over affective primacy. However, Nakashima's temporal judgment method has a number of potentially serious shortcomings when compared to standard TOJ and SJ tasks (see Reisenzein & Franikowski, 2021). Furthermore, Nakashima did not use temporal judgments to time object recognition for affective pictures, the stimuli most often used in the subliminal perception studies and in the recent RT studies on affective vs. semantic primacy (Nummenmaa et al. 2010)." was corrected to: "The use of temporal judgments in affect research was again pioneered by Nakashima (1909a, Exp. 1–3), who in three additional studies used a self-developed procedure called the 'direct reaction method' to time the feelings evoked by colors, geometric figures and affective pictures of natural objects. The participants were presented with each affective stimulus for different durations, and were asked to report whether they experienced the feelings evoked by the stimulus while it was still visible, or only later. This method can be regarded as a variant of the TOJ procedure in which the probe stimulus consists of the offset of the target stimulus. While Nakashima (1909a) interpreted the results of these studies as also supporting semantic over affective primacy, he actually did not time object recognition with the 'direct reaction' method. Furthermore, this method has potentially serious shortcomings when compared to standard TOJ and SJ tasks (see Reisenzein & Franikowski, 2021)."

- Section **Latencies of Object Recognition and Affect** of the Results section for Experiment 1, paragraph 5: "picture onset position" was corrected to: "picture onset time"

The final article is available, upon publication, at: https://doi.org/10.1037/emo0000991.supp

Perceptual Latencies of Object Recognition and Affect Measured With the Rotating Spot Method: Chronometric Evidence for Semantic Primacy

Philipp Franikowski, Lara-Sophie Kriegeskorte, and Rainer Reisenzein

Supplemental Materials S1: Supplementary Analyses

Experiment 1

Preliminary ANOVA

A preliminary ANOVA was conducted to explore possible effects of the experimental control factors picture onset time (picture presented after 107, 214, 321 or 428 ms), valence order (pleasant or unpleasant pictures judged first), and valence (pleasant or unpleasant pictures). This ANOVA included the control factors plus the factor of main interest, task (object recognition versus affect). All factors with the exception of valence order were within-subject. The degrees of freedom of the F-test of picture onset time were Huynh-Feldt corrected to protect against violations of sphericity, and the p-values of all effects involving the control variables were Holm-corrected (see Cramer et al., 2016).

The ANOVA found no significant main or interaction effects for valence order and picture valence $F \leq 3.75$, $p \geq .779$, but detected a significant main effect of picture onset time, $F(1.95, 31.2) = 42.42$, $p < .001$, generalized $\eta^2 = .06$, as well as a significant interaction between task and picture onset time, $F(2.95, 47.22) = 9.05$, $p < .001$, $\eta^2_g = .01$. In addition, the main effect of task was significant $F(1, 16) = 46.07$, $p < .001$, $\eta^2_g = .34$.

Experiment 2

Affect Ratings

The mean ratings of the positive pictures ranged from 6.4 to 8.5 ($M = 7.4$, $SD = 0.7$) if the pictures were not blurred, and from 5.0 to 8.3 ($M = 6.7$, $SD = 1.2$) if they were blurred; the mean unpleasantness ratings of the negative pictures ranged from 5.2 to 8.3 ($M = 6.8$, $SD = 1.0$) if the pictures were sharp, and from 3.5 to 6.7 ($M = 5.5$, $SD = 0.9$) if they were blurred. For individual participants, the average pleasantness rating of the positive pictures ranged from 3.7 to 10.0 ($M = 7.4$, $SD = 1.4$) if they were sharp and from 3.8 to 10.0 ($M = 6.7$, $SD = 1.6$) if they were blurred, and the average unpleasantness rating of negative pictures ranged from 3.5 to 9.7 ($M = 6.8$, $SD = 1.8$) if the pictures were sharp and from 2.5 to 9.2 ($M = 5.5$, $SD = 1.9$) if they were blurred. Taken together, the affect ratings indicate that most pictures induced from moderate to strong pleasant or unpleasant feelings in all participants.

Preliminary ANOVA

As in Experiment 1, a preliminary ANOVA was conducted to explore possible effects of the experimental control factors. In Experiment 2, this ANOVA included the factors of main interest, task (object recognition versus affect onset) and blur status (pictures blurred or not blurred), plus the control variables picture onset time, valence order (positive or negative pictures judged first), task order (object recognition or affect detection first), picture valence (pleasant or unpleasant pictures), and blur group (picture set A or B blurred). Task type, picture onset time, blur status and picture valence were within-subject factors, whereas valence order, task order and blur group were between-subject factors. Huynh-Feldt corrections of the degree of freedoms were applied when appropriate; the p-values for the control factors were Holm-corrected. This ANOVA revealed no significant main effects or interactions for the control variables valence order (second ANOVA: $F \leq 4.45$, $p > .999$), but found a main effect of picture onset time, $F(1.94, 58.07) = 72.89$, $p < .001$, $\eta^2_g = .08$, as in Experiment 1. The main effect of task was also significant, $F(1, 30) = 33.73$, $p < .001$, $\eta^2_g = .16$, as was the factor blur status, $F(1, 30) = 36.28$, $p < .001$, $\eta^2_g = .02$.

Appendix C

Publication C

On the latency of object recognition and affect: Evidence from speeded reaction time tasks

Franikowski, P., & Reisenzein, R. (2022).
Emotion. Advance online publication.
https://doi.org/10.1037/emo0001092

The final article is available, upon publication, at: https://doi.org/10.1037/emo0001092

On the Latency of Object Recognition and Affect: Evidence from Speeded Reaction Time Tasks

Philipp Franikowski and Rainer Reisenzein
University of Greifswald

Abstract

According to the semantic primacy hypothesis of emotion generation, stimuli must be semantically categorized to evoke emotions. This hypothesis was tested in two speeded reaction time experiments that also explored the processes underlying valence judgments. Participants viewed pleasant and unpleasant pictures. In different blocks of trials, they pressed a key as soon as they experienced the feeling evoked by a picture, recognized the depicted object, or detected the valence (pleasant/unpleasant) of the picture. Object recognition was significantly earlier than affect onset, and the two latencies were positively correlated. The latency of valence detection was in between the latencies of object recognition and affect, and correlated with both. Experiment 2 additionally found that blurring the pictures delayed the onset of affect, and that this effect was partially mediated by delayed object recognition. In contrast, false-coloring the pictures was found to delay affect mainly by reducing its intensity. Coloring also delayed valence judgments of pleasant pictures. The findings provide further chronometric support for the semantic primacy hypothesis of emotion generation and shed light on the processes underlying valence judgments.

Keywords: semantic vs. affective primacy; mental chronometrics; latency of object recognition; latency of affective experience; speeded reaction time task

The semantic primacy hypothesis posits that perceived objects must be recognized (i.e., subsumed under a concept, or classified as an instance of a semantic category) to evoke emotional reactions (e.g., Lazarus, 1982; Ortony et al., 1988; Storbeck et al., 2006). For example, to experience a pleasant feeling when seeing a picture of a cute cat in a magazine (Myrick, 2015), one must recognize the depicted object as being a cat. According to the proponents of semantic primacy, the recognition of the object is needed because it constitutes the basis of the evaluation of the object, which is regarded as the immediate cause of affect (e.g., Lazarus, 1982). That is, objects are evaluated as positive or negative, and consequently evoke affect, *because* they are recognized as being of a particular kind (e.g., a cat; rotten food).

The semantic primacy hypothesis has been vigorously contested by "noncognitive" emotion theorists (e.g., James, 1890; and more recently, LeDoux, 1998; Öhman

---

Author note

Philipp Franikowski, https://orcid.org/0000-0002-2109-7388, Rainer Reisenzein, https://orcid.org/0000-0002-6091-108X, Institute of Psychology, University of Greifswald

Correspondence concerning this article should be addressed to Philipp Franikowski, Institute of Psychology, University of Greifswald, Franz-Mehring-Str. 47, 17489 Greifswald. E-mail: philipp.franikowski@uni-greifswald.de

All analyzed data and R scripts are available on https://osf.io/2prm4/.
This study was not preregistered.

& Mineka, 2001; Zajonc, 1980). These theorists have argued that at least some emotions can be – and perhaps even, regularly are (Zajonc, 1980) – evoked by stimuli via a "noncognitive" route that circumvents object recognition and recognition-based evaluations. Because of its potentially far-reaching implications, this alternative hypothesis of emotion generation, the (causal version of the) affective primacy hypothesis, has attracted much interest not only in psychology, but also in neighboring sciences (see Leys, 2017). The most plausible explication of the affective primacy hypothesis, in our view, is that "noncognitively" caused emotions are elicited by *nonconceptual* or *preconceptual representations* (see e.g., Crane, 1992; Marr, 1982) of the eliciting objects. In the visual case, these could be image-like representations of the prototypical shapes, textures and color patterns of the emotion-evoking objects (e.g., Öhman & Mineka, 2001).

Empirical Evidence

Seemingly supporting the affective primacy hypothesis, numerous studies reported that subliminally presented affective stimuli evoke reactions that suggest the occurrence of an emotion, such as peripheral physiological reactions (e.g., Öhman & Mineka, 2001), facial expressions (e.g., Dimberg et al., 2000) and increased amygdala activity (e.g., Morris et al., 1998) (for a review, see Lähteenmäki et al., 2015).

However, these findings are inconclusive for at least two reasons. First, even if the findings are accepted at face value, they do not prove affective primacy, because the semantic recognition plus evaluation of objects could have occurred at an unconscious level of processing (e.g., Lazarus, 1995; Storbeck et al., 2006). Second, more recent research suggests that the presentation times used in the subliminal perception studies were not short enough to prevent at least partial awareness of the stimuli in a subset of participants or trials (e.g., Lähteenmäki et al., 2015; Pessoa et al., 2005). Recent experiments using a more sensitive psychometric threshold detection procedure and more sensitive measures of object recognition found that affective pictures elicited physiological reactions (e.g., Peira et al., 2012; see also Codispoti et al., 2009), and could be reliably judged as pleasant or unpleasant (Lähteenmäki et al., 2015) or as evoking pleasant or unpleasant feelings (Peira et al., 2012), only at presentation times at which the depicted objects were at least partly recognized. These findings suggest that the elicitation of emotions may not just require the *recognition* of the eliciting objects, but even their *conscious* recognition (for additional evidence, see Codispoti et al., 2021; Hedger et al., 2015; Schupp et al., 2008).

This conclusion is supported by findings obtained with a second, albeit much less-used experimental paradigm, in which the latency of affective feelings evoked by objects is compared to the time needed to recognize the objects (e.g., Franikowski et al., in press; Nakashima, 1909a, 1909b; Nummenmaa et al., 2010; Reisenzein & Franikowski, 2021). The reasoning behind these chronometric studies is as follows: Causes must precede their effects (Mackie, 1974); therefore, if the semantic primacy hypothesis is correct, the latency of object recognition must be shorter than the latency of affect. Conversely, if the latency of object recognition is longer than the latency of affect onset, the semantic primacy hypothesis is refuted, suggesting that – in the absence of an alternative – the affective primacy hypothesis is true.

There are two traditional methods for measuring the latency of conscious mental events (Neumann & Niepel, 2004): Speeded reaction time (sRT) tasks, and temporal judgment (TJ) tasks (e.g., Jaśkowski, 2014). Both chronometric methods were first applied to affect by a student of Edward B. Titchener, Taiko Nakashima (1909a; 1909b), who tested an early version of the affective primacy hypothesis proposed by Wundt (1896). In three of his studies, Nakashima (1909a, 1909b) used sRT tasks to measure how long it takes to report the occurrence of pleasant and unpleasant feelings evoked by simple stimuli (colors, tones, touch), and the recognition of the stimuli. In conflict with the affective primacy hypothesis, but consistent with semantic primacy, the recognition judgments were found to take less time than the judgments of affect onset. Subsequent sRT experiments by Wells (1925, 1929) partly confirmed these findings.[1]

Following these pioneering studies, chronometric research on the experience of affect and object recognition was abandoned until recently, when Nummenmaa et al. (2010) compared the latency of semantic judgments of pictures (e.g., does the picture contain an animal or not?) to "affective" judgments (judgments of stimulus valence: is the picture [un]pleasant or not?). Across five experiments, the semantic judgments were found to take consistently less time than the affective judgments. This confirms the above-reported findings of Nakashima (1909a; 1909b) if one assumes that the valence judgments were based, at least in substantial part, on the feelings evoked by the stimuli. This important issue is discussed in more detail below.

In additional studies, Nakashima (1909a, Studies 1–3) measured the latency of affect using a self-devised temporal judgment method that can be reconstructed as a variant of the temporal order judgment method (see Reisenzein & Franikowski, 2021). However, the speed of object recognition was not measured in these studies, and therefore no direct comparison of the latencies of cognition and affect could be made. This line of research was recently taken up again by Reisenzein and Franikowski (2021) and Franikowski et al. (in press). Reisenzein and Franikowski (2021) used temporal order judgments (TOJ) and simultaneity judgments (SJ) to time object recognition and affect onset. The participants watched pleasant and unpleasant pictures that were preceded or followed, at variable intervals, by a probe stimulus (a blue bar). In one block of trials, the participants judged when they noticed the pleasant or unpleasant feeling evoked by the picture: before or after the probe (TOJ), or simultaneously with the probe or not (SJ). In another block of trials, they similarly judged when they recognized the objects shown in the picture. The latencies of object recognition and affect were estimated from these data by computing the point of subjective simultaneity (PSS), the delay between target picture and probe stimulus at which the probe appeared to be simultaneous with object recognition or affect, respectively. In agreement with Nakashima's (1909a) conclusions, the PSS of affective experience occurred significantly later than the PSS of conscious object recognition.

In a follow-up study, Franikowski et al. (in press) used the rotating spot (or rotating clock-hand) method (e.g., Libet et al., 1983; Pockett & Miller, 2007) to time the occurrence of object recognition and affect. Participants watched a rotating spot symbolizing a clock-hand that quickly revolved around affective pictures. In different blocks of trials, the participants memorized the position of the spot at the time when they recognized the object shown in the picture, or experienced the affect

[1]Wells (1925, 1929) was mainly interested in the latency of affect and therefore did not include precisely matched object-recognition judgments. However, the measured sRT affect latencies (about 800 ms) were close to those reported by Nakashima (1909a; 1909b).

evoked by the picture. Again, objects were recognized significantly faster than affect was experienced.

The results of the described chronometric studies are also in fair agreement with findings on the latency of the late positive potential (LPP), the most reliably affect-sensitive component of the ERP (e.g., Codispoti et al., 2006; Cuthbert et al., 2000; see Hajcak & Foti, 2020, for a recent review). The LPP is a sustained positive deflection in the ERP waveform that begins 300–400 ms after stimulus onset and is larger in response to pleasant and unpleasant than neutral stimuli. According to a widespread view, the affective modulation of the LPP indicates the attentive processing of motivationally significant stimuli (Cuthbert et al, 2000; Hajcak & Foti, 2020). Although the LPP is thus not a direct neurophysiological marker of pleasant and unpleasant feelings, it is phenomenologically plausible, as well as consistent with several theoretical models of emotion (e.g., Bradley, 2009; Reisenzein, 2009), to assume that these feelings set in close to the allocation of attention to the emotional stimuli, and hence to the onset of the LPP. In contrast, sRT and ERP studies suggest that object recognition is already possible within 200 ms (e.g., Fabre-Thorpe, 2011; Johnson, & Olshausen, 2003), and thus 100–200 ms before the affective modulation of the LPP. This finding is in good agreement with the results of the temporal judgment studies summarized above, which found an average delay between object recognition and affect onset of 120–150 ms (Franikowski et al., in press; Reisenzein & Franikowski, 2021).

In this article, we report two new sRT studies that compared the latencies of object recognition and affective experience. The studies had three goals. The first goal was to replicate the basic findings of our temporal judgment studies (Franikowski et al., in press; Reisenzein & Franikowski, 2021) with sRT tasks. The second goal, pursued in Experiment 2, was to replicate and extend an experimental test of the semantic primacy hypothesis devised by Franikowski et al. (in press). The third goal was to compare the latencies of object recognition and affect to the time needed to detect stimulus valence (Nummenmaa et al., 2010). These goals are now explained in more detail.

Cross-Method Replication

The first goal of the experiments was to replicate our previous findings on the latencies of object recognition and affect with sRT measurements. To optimize the comparison, we used the same stimuli, presentation conditions, and – most important – the same kinds of judgments as in the temporal judgment studies. The cross-method replication was motivated by the assumption that each latency measurement method has its weaknesses, but different methods have different weaknesses (see Linares & Holcombe, 2014). Therefore, if the cross-method replication attempt is successful, the conclusions drawn from the results obtained with any one method are strengthened.

A major advantage of temporal judgments, compared to sRT measurements, is that they provide a purer estimate of perceptual latencies, because they do not include the time to decide on, and execute, the motor response (Schneider & Bavelier, 1983). An advantage of sRT measurements is that they avoid possible interference artifacts caused by the dual-task nature of TJ tasks (see Franikowski et al., in press, for a discussion). Additional differences between sRT tasks and temporal judgments are discussed in Cardoso-Leite and Gorea (2010), Jaśkowski (2014), and Miller and Schwarz (2006).

The two main findings of the temporal judgment studies (Franikowski et al., in press; Reisenzein et al., 2021) that we hoped to replicate with the sRT method were: (a) the latency of object recognition is shorter than the latency of affect onset; (b) there is a positive correlation between the two latencies.

Extended Experimental Test of the Semantic Primacy Hypothesis

A limitation of most previous tests of the semantic primacy hypothesis in the latency measurement paradigm is that these tests were nonexperimental; that is, object recognition was not experimentally manipulated. To overcome this limitation, Franikowski et al. (in press, Exp. 2) delayed object recognition by blurring the affective pictures (see De Cesarei & Codispoti, 2011). As predicted by the semantic primacy hypothesis, this manipulation also delayed the onset of affect, and the effect of blurring on affect latency was found to be partly mediated by the delay in object recognition. Experiment 2 was an extended replication of this study in the sRT paradigm. The extension consisted of the inclusion of a second experimental manipulation, false-coloring the pictures. The aim of this manipulation was to clarify the influence of affect intensity on affect latency.

Research on the latencies of sensory qualities has consistently found that high-intensity sensations (e.g., loud tones) are detected earlier than low-intensity sensations (e.g., low tones; see Cardoso-Leite & Gorea, 2010; Jaśkowski, 2014). Given that feelings of pleasure and displeasure also have an intensity, and are in fact often theorized as being similar to sensations (see Reisenzein, 2012), a parallel intensity effect can be expected for affect. Suggestive evidence for this hypothesis was obtained in the study by Franikowski et al. (in press), who found that picture blurring, apart from delaying affect by delaying object recognition, also had a direct effect on affect latency. Because blurring also reduced the intensity of experienced feelings, this latter effect could have been mediated by reduced affect intensity. However, the findings were inconclusive (see Franikowski et al., in press).

To clarify the effects of affect intensity on affect latency, Experiment 2 included a second image manipulation that has been found to strongly reduce the intensity of picture-evoked affect, false-coloring (Junge & Reisenzein, 2013, Exp. 2). The reason may be that false-coloring pictures reduces the similarity of the depicted objects to affect-evoking prototypes stored in memory (Junge & Reisenzein, 2013). In contrast to blurring, however, false-coloring was expected to have only a small, if any, effect on the latency of object recognition. The reason is that false colors leave the coarse and fine contours of the depicted objects largely intact, and these features, which carry information about the objects' shape, are of main importance for object recognition (e.g., Morrison & Schyns, 2001). Color seems to influence the speed of object recognition only under special conditions, in particular if the colors are diagnostic for the object in question (see Bramão et al., 2011). Even then, the temporal facilitation of object recognition by color is typically small (Weiß & Scharlau, 2010).

In sum, Experiment 2 extended the previous experimental test of the semantic primacy hypothesis by comparing the effects of blurring, a manipulation that has been found to influence the latency of affect by delaying object recognition, to that of false-coloring, a manipulation expected to influence affect latency by reducing affect intensity.

Affect Versus Stimulus Valence

The third, more exploratory goal of our experiments was to compare the latency of experienced affect with the latency of detecting stimulus valence. In the chronometric studies of Nakashima (1909a; 1909b), Reisenzein and Franikowski (2021) and Franikowski et al. (in press), the participants were asked to indicate *when they noticed the pleasant or unpleasant feeling* evoked by an affective stimulus. This question unambiguously asks for the occurrence of an affective reaction. In contrast, Nummenmaa et al. (2010), as well as Wells (1925, 1929) and in one study Nakashima (1909b, Study 3), asked their participants to indicate *whether or not the stimulus was pleasant or unpleasant*. This question does not directly refer to the person's feelings but asks for a property of the stimulus, its affective valence (Russell, 2003).[2] Nevertheless, valence judgments of stimuli can be used as proxies for affect reports if one is willing to assume that they are based on the feelings evoked by the stimuli (e.g., that a picture is judged as pleasant because it causes a pleasant feeling when looking at it; see also, Clore & Byrne, 1974). This assumption was indeed made by Nakashima (1909b, Study 3) and Wells (1925, 1929), and it has also been made in some recent studies that used valence judgments to measure picture-evoked affect (e.g., Kurdi et al., 2017). However, as already pointed out by Wells (1925), this assumption is not necessarily valid: valence judgments can be the result of purely cognitive processes.

Valence judgments are a special form of evaluative judgments that attribute a specific evaluative property, hedonic goodness or badness, to objects. According to a common analysis (e.g., Meinong, 1894; Russell, 2003), the affective valence (pleasantness or unpleasantness) of a stimulus or object is a dispositional property of the object, its capacity or tendency to evoke pleasant or unpleasant feelings in the perceiver under appropriate circumstances. On this analysis, noticing the feelings evoked by an object is the *epistemically primary* way of learning the valence of the object (Meinong, 1894; see also, Clore & Byrne, 1974), and it may also be the *usual* way of how the pleasantness of objects is determined, particularly when the objects are novel. However, once the hedonic valence of an object is known (e.g., "The tulips on my table are pleasant"), a valence judgment can be made more easily, as well as more quickly, by retrieving the stored valence information from the memory schema of the object (see e.g., Fazio, 2007; Forgas, 1995; Itkes et al., 2017). Alternatively, an object's affective value could be estimated by inferring it from the stored valence of the category to which the object belongs (Fiske, 1982; e.g., "Tulips are generally pleasant, therefore the tulips on my table are pleasant as well"), or from the stored valences of object properties or components (e.g., "the colors of the tulips are pleasant", "the leaves have a pleasant shape"; e.g., Fishbein, 1963; Kaplan & Anderson, 1973).[3]

[2]Note that Russell (2003) and other authors (e.g., Barrett, 1996) also use the term "valence" to denote the hedonic tone (pleasure vs. displeasure) of emotional experiences. These two meanings of "valence" (stimulus valence vs. the valence of emotional experiences) are related as follows: A stimulus is pleasant (has positive valence) for a person if it evokes pleasant feelings of positive hedonic valence) in that person under appropriate circumstances, and analogously for negative valence. In addition, "valence" is often used more generally to denote the subjective value of an object or event, regardless of the basis of the value (e.g., Lewin et al., 1944).

[3]Note also that some participants could interpret a question asking for the pleasantness of a picture in a liberal or metaphorical way, as asking whether the depicted objects are in some way good or bad (Osgood et al., 1957).

To which degree the valence judgments in previous chronometric studies were based on the pleasant and unpleasant feelings evoked by the judged stimuli versus on stored valence information is not known. The above considerations suggest, however, that judgments of valence are more likely feeling-based if the judged object is novel (as also proposed by Forgas, 1995; Schwarz & Clore, 2007). In contrast, if the object is well-known (e.g., because it has been presented repeatedly in an experiment, or because fits a well-known prototype), valence judgments may be made by retrieving stored object valences. Some evidence for this hypothesis was reported by Itkes et al. (2017), who found that the intensity of subjective feelings, as well as facial EMG reactions and heart rate changes evoked by affective pictures, declined after many repetitions of the stimuli, whereas (knowledge-focused) valence judgments showed no habituation. Arguably consistent with these findings, other studies found rapid habituation of peripheral physiological reactions, as well as the overall magnitude of the LPP, to repetitions of affective pictures, whereas the affective modulation effect (the difference in the LPP between affective and neutral pictures) remained largely intact (Codispoti et al., 2006; Ferrari et al., 2020; Micucci et al. 2020). Note, however, that valence judgments could be memory-based even if the stimuli continue to evoke strong affective reactions, for example if fast responding is required.

The two described hypotheses about the processes underlying valence judgments make relatively straightforward predictions for the latencies of these judgments, compared to the latencies of object recognition and affect. (1) Feeling-based valence judgments of an object can only be made after the feeling evoked by the object has been registered. They should therefore take about as much time as affect reports, and their latency should correlate with the latency of affect. (2) Memory-based valence judgments of an object can be made as soon as the object schema has been activated by a stimulus. They should therefore take about the same time as object recognition judgments, and the two latencies should be correlated. (3) If the valence judgments of a set of n objects are partly feeling-based and partly memory-based (i.e. k are feeling-based and $n - k$ are memory-based), then the average latency of valence detection of the n objects should lie in between the latencies of object recognition and affect, and should correlate with both latencies. In our view, this is the most likely scenario for valence judgments of affective pictures in typical experiments, particularly if the pictures are presented repeatedly.

The assumption that valence judgments in sRT tasks are, at least in substantial part, memory-based can explain why the cognition-affect delay found in the RT studies of Nummenmaa et al. (2010) was only about half the size of the delay found in the temporal judgment studies of Reisenzein and Franikowski (2021) and Franikowski et al. (in press), despite the fact that latency differences measured with sRT tasks are typically larger than those measured with TJs (Cardoso-Leite & Gorea, 2010). However, to test this hypothesis in a stringent manner, it is necessary to compare the latencies of affect and valence detection for the same stimuli under identical viewing conditions. This was done in our experiments. To increase the probability of memory-based judgments, and thus the chances of finding the predicted effects, the valence judgment task was placed at the end of the experiment. This gave the participants repeated (altogether, five) opportunities to experience the pleasure or displeasure evoked by each picture, and to store this information in the memory schemas of the depicted objects.

Hypotheses

The cross-method-replication aspect of the present studies concerns primarily (see H9 and H10) three predictions derived from the semantic primacy hypothesis.

H1: The latency of object recognition is shorter than the latency of affect (tested in Experiments 1 and 2).

H2: There is a positive correlation between the latencies of object recognition and affect (Exp. 1 and 2).

H3: (a) Blurring the affective pictures causes a delay of object recognition and affect onset; (b) the effect of blurring on affect onset is mediated, at least in part, by delayed object recognition. (Exp. 2).

H1 and H2 were already tested and supported in the TJ paradigms used by Franikowski et al. (in press) and Reisenzein and Franikowski (2021). H3 was tested, and also supported, in Franikowski et al. (in press, Exp. 2). Here we sought to replicate H1–H3 with sRT tasks, using the same stimuli and presentation conditions.

The first new hypothesis, tested in Experiment 2, concerns the effects of false-coloring.

H4: (a) False-coloring the pictures increases the latency of affect, but has no or only a small effect on the latency of object recognition. (b) The effect of false-coloring on affect latency is mediated by reduced affect intensity.

Four additional new hypotheses, three of them tested in both experiments and the fourth in Experiment 2, concern the latency of valence detection.

H5: The latency of valence detection is longer than the latency of object recognition.

H6: The latency of valence detection is shorter than the latency of affect.

H7: The latency of valence detection correlates positively with (a) the latency of affect and (b) the latency of object recognition.

H8: Blurring and false-coloring the pictures both delay valence judgments.

On the data level, H5 simply asserts that the findings of Nummenmaa et al. (2010) will be replicated with our stimulus set and presentation mode (at least 2000 ms stimulus presentation, rather than 30 ms). On the level of underlying processes, H5 is derived from the assumption that valence judgments can be either feeling-based or memory-based, and that both strategies are used for different pictures. H6, H7, and H8 can be derived from the same premises. Specifically, H8 can be derived as follows: Affect-based valence judgments are delayed by all factors that delay affect onset; these factors include (we assume) blurring and coloring. Memory-based valence judgments are delayed by blurring because (we assume) blurring delays the activation of the object schema in which valence information is stored, i.e., the time until the schema information becomes conscious.

In addition to testing these hypotheses, we expected to replicate the findings of Franikowski et al. (in press) that color recognition judgments are faster than object recognition and affect onset judgments (H9), and that judgments of stimulus appearance are faster than color recognition judgments (H10).

The ten hypotheses can be compactly summarized as follows, using "$x \prec y$" for "the latency of x is shorter than the latency of y": (1) H1, H5, H6, H9 and H10 together imply: stimulus appearance detection $\prec$ color recognition $\prec$ object recognition $\prec$ valence detection $\prec$ affect onset. (2) The correlational hypotheses H2, H7a and H7b together imply that the latencies of object recognition, affect, and valence detection are positively intercorrelated. (3) The remaining, causal hypotheses assert: (H3) picture blurring delays affect, and the blurring→affect latency effect is, at least in part, mediated by delayed object recognition; (H4) picture coloring also delays affect, but the coloring→affect latency effect is mediated by reduced affect intensity rather than delayed object recognition; and (H8) both blurring and coloring delay valence judgments.

Ethics statement

All experiments were designed and conducted in accordance with the Code of Ethics of the World Medical Association ("World Medical Association Declaration of Helsinki," 2013). All participants in the two experiments gave their informed consent to participate prior to the start of the experiments. They were informed that they would see pleasant and unpleasant pictures and that they could terminate their participation at any time without negative consequences. Ethical review and approval were not required for these studies in accordance with the German legislation and with institutional requirements. In the following description of our experiments, we report all measures, manipulations, and exclusions (if any). This study was not preregistered. All analyzed data and R scripts are available on https://osf.io/2prm4/

Because Experiment 2 was conducted during the Sars-CoV-2 epidemic in late 2020, strict measures were taken to minimize the risk of infection: Only one subject participated at a time; the experimenter was not present in the room during the experiment; participants and experimenter wore FFP2 face masks; their interaction was limited to a brief introduction at the beginning and participant remuneration at the end of the experiment; the room was aired out between sessions for half an hour; and the participant's work place was disinfected after each session.

Experiment 1

Method

Participants

The participants were 20 undergraduates, 14 females and 6 males (age M = 25.4 years, SD = 7.8). They were randomly assigned to one of the four between-subject conditions of the experiment. Two participants were excluded from the data analyses because of very low affect ratings (see Results), reducing the final sample size to 18. The participants were given course credit or received a small fee for their participation.

The minimum sample size required for detecting the focal latency difference between object recognition and affect was estimated from the studies of Reisenzein and Franikowski (2021, Exp. 1–2), and Franikowski et al. (in press, Exp. 1–2). The effect size of the cognition-affect lag obtained in these studies ranged from $\eta_g^2 = .24$ to .36 (Mdn = .26), and the correlation between the two latencies ranged from .19 to .61 (Mdn = .31). To detect an effect of $\eta_g^2 = .26$ with a power of .80, assuming a correlation of .31 between the latencies, a sample size of 10 would be sufficient (Faul et al., 2007). The actual sample size of $N = 18$ used in the data analysis detects this effect with a power of .98.

Stimuli and apparatus

We used the same stimuli as in the TOJ and SJ experiments of Reisenzein and Franikowski (2021). For object recognition, affect, and valence detection, the stimuli consisted of 12 pleasant and 12 unpleasant (disgusting) pictures; for color recognition, they were six uniformly colored pictures (blue, green, red, yellow, black, and white). All pictures were 300 pixels wide and 360 pixels high. The affective pictures were collected from various free sources on the internet; if necessary, they were cut and resized. The pleasant pictures included pictures of a laughing child, a bunny, a kitten, a panda bear, a sunflower field and an ice cream cup; the disgusting pictures showed among others a purulent finger, vomit, a snake pit, maggots, a moldy piece of bread, and an overflowing ashtray. Franikowski et al. (in press) and Junge and Reisenzein (2013) found that these pictures elicited from medium to strong pleasant and unpleasant feelings in most participants. Previous research has found that comparable pleasant and unpleasant pictures also tend to evoke weak physiological reactions and facial expressions (activation of the zygomaticus vs. corrugator muscle) in viewers (e.g., Codispoti et al., 2009; Lang et al., 1993), as well as a modulation of the LPP (e.g. Bradley et al., 2000; Codispoti et al., 2009).

The pictures were presented at the center of a monitor (75 Hz refresh rate) with a resolution of 1024 × 1280 pixels at normal viewing distance. The experiment was programmed in DMDX (Forster & Forster, 2003). Responses were collected via keyboard input.

Procedure

The experiment was conducted in a computer lab with one to four subjects participating at a time. Workplaces were separated by room dividers. All participants performed five tasks in the following sequence: Affect ratings, color recognition, object recognition followed by affect or vice versa, and valence judgments. The experiment lasted about half an hour.

Affect ratings. To familiarize the participants with the pictures and to verify that they evoked at least moderate degrees of pleasant or unpleasant affect in each participant, the participants first rated the affective pictures for the degree of pleasure or displeasure they evoked ("How pleasant [unpleasant] do you experience this picture to be?"). Pleasant and unpleasant pictures were rated in separate blocks and were presented in an individualized random order within each block. Answers were given on 11-point rating scales ranging from 0 (not at all pleasant/unpleasant) to 10 (extremely pleasant/unpleasant; the German words were angenehm/unangenehm). Although the negative pictures were all disgusting in nature, we used the term "unpleasant" rather than "disgusting" as the scale label to make the two affect ratings as similar as possible.

Reaction Time Tasks. Four types of latency were measured, the latency of color recognition, object recognition, affect onset, and valence detection. The first three tasks were simple RT tasks in which participants pressed a designated key (the letter "B" on the German computer keyboard) as quickly as possible when they recognized the color of a uniformly colored picture, recognized the object shown in an affective picture, or experienced the pleasant or unpleasant feeling evoked by this picture. The fifth task was a 1AFC (one-alternative forced choice) RT task in which the participants categorized the affective pictures as pleasant or unpleasant as fast as possible by pressing one of two labeled keys (the keys immediately to the left and right of the spacebar).

In the color recognition trials, each color picture (i.e., red, green, blue, yellow, black, or white) was presented four times, resulting in 24 trials. One purpose of the color recognition task was to familiarize the participants with the general procedure of the subsequent object recognition and affect tasks; the second was to provide additional data on the validity of the RT measurements, as color recognition was expected to be faster than object recognition. In the object recognition task, the participants viewed, in separate blocks, the pleasant and unpleasant pictures and were asked to press the response key as soon as they recognized which object or objects were shown (e.g., a panda bear, a sunflower field, an ashtray). Each picture was presented twice, resulting in 24 trials per valence block or 48 trials in total. In the affect detection task, the participants saw the same pictures, but pressed the response key as soon as they noticed the pleasant or unpleasant feeling evoked by the picture. Again, each picture was presented twice, resulting in 24 trials per valence block and 48 trials in total.

In the valence detection task, the pleasant and unpleasant pictures were shown in a randomly intermixed sequence and the participants pressed one of two response keys as soon as they recognized that the picture was pleasant or unpleasant, respectively. The 1AFC task was chosen because it allowed us to stay as close as possible to the valence detection task used in the sRT experiments of Nummenmaa et al. (2010) without having to introduce new (neutral) stimuli at this point.[4] Also, because we wanted to use stimulus presentation times comparable to those used in our TJ experiments, the pictures were shown for a much longer time (at least 2000 ms) than in the studies of Nummenmaa et al. (30 ms). Analogous to the other sRT tasks, the 1AFC task was repeated twice to assess retest reliability. Hence, it comprised 48 trials in total.

Trial Structure. Each trial of the sRT tasks began with a single-line message that reminded the participant of the required judgment (e.g., "Please press the 'B' key as soon as you recognize the object or objects shown"). After 2000 ms, a gray picture was presented at the center of the screen for a random duration ranging from 400 to 907 ms, to focus attention on the location of the upcoming target picture. Subsequently, the target picture was presented and remained on the screen for at least 2000 ms or until the participant responded (however, most responses were given within the 2000 ms window). All responses and RTs were collected by the computer program.

Control of Sequence Effects. To facilitate the detection of affect onset, the

[4]Nummenmaa et al. (2010) used a 2AFC version of the valence detection task, that is, the participants saw two pictures in each trial, one pleasant or unpleasant, and the other neutral, and indicated the position of the valenced picture.

positive and negative pictures were judged in separate blocks, as this allowed the participants to focus on a feeling of known valence (pleasant or unpleasant).[5] To control for possible sequence effects, the order of the pleasant and the unpleasant picture set and the sequence of the object recognition and affect tasks within the valence blocks was counterbalanced across subjects. This meant that, following the color recognition task, participants were randomly assigned to four groups resulting from the combination of valence order and task order. Within each block of trials, the order of the pictures was individually randomized to counteract possible effects of item sequence.

Data analysis. Because reaction times usually have right-skewed distributions and are typically contaminated with long-latency outliers caused by lapses of attention and other disturbances, a robust location measure such as the median or the trimmed mean is often recommended to represent the central tendency (e.g., Whelan, 2008; Wilcox, 2016). Because one goal of our studies was to compare our findings to those obtained by Nummenmaa et al. (2010), we followed these authors in using the median. That is, the individual RTs were first aggregated to medians for each participant for each experimental condition of interest (see below) and the median RTs were then processed further by mixed-design analysis of variance (ANOVA) and *t*-tests. The data analyses were conducted in *R* (R Core Team, 2017). The ANOVAs were performed using the add-on package *afex* (Singmann et al., 2020). The *p*-values of predicted effects were uncorrected, whereas all other effects were Holm-corrected (Holm, 1979) to account for error accumulation in multiway ANOVAs (Cramer et al., 2016). Generalized η^2 (η^2_g, Bakeman, 2005) was used as the effect size measure for the ANOVAs and *t*-tests. The correlational hypotheses were tested by computing between-subjects correlations between the median latencies of individuals.

Results

Two of the 20 participants were excluded from the data analyses because they gave zero or very low "unpleasant" ratings to most negative (disgust) pictures, suggesting that these pictures failed to evoke negative feelings in them. The final sample therefore comprised 18 participants, 13 females and 5 males (age M = 25.5 years, SD = 8.2).

Affect Ratings

For the 18 participants included into the analyses, the mean pleasantness rating of the positive pictures ranged from 6.6 to 8.4 (M = 7.4, SD = 0.6); the mean unpleasantness rating of the negative pictures ranged from 4.4 to 8.4 (M = 6.8, SD

[5] A possible objection to the block-wise presentation of same-valence pictures is that the picture-evoked affect might accumulate across trials, which could make the detection of the onset of affect evoked by each new picture against the increasing baseline increasingly difficult, resulting in an artificial RT increase. However, a trend analysis of the median reaction times across the trials of the affect task yielded no evidence for this hypothesis (b = -2.57, 95% CI = [-6.64, 1.50], p = .213). In addition, trend analyses of the mean ratings in the rating blocks for pleasant and unpleasant pictures across trials yielded no evidence for an accumulation of affect (b = -0.02, 95% CI = [-0.09, 0.04], p = .369). Also speaking against an accumulation of affect, the pictures in the present studies were separated by an intertrial interval between 2400 to 2907 ms, during which time the response instruction (2000 ms) and the attention-focusing picture (400–907 ms) were shown.

= 1.2). For individual participants, the average rating of the 12 pleasant pictures ranged from 4.6 to 8.7 (M = 7.4, SD = 1.2) and that for the 12 unpleasant pictures, from 4.5 to 9.7 (M = 6.8, SD = 1.3). These findings closely replicate those of Reisenzein and Franikowski (2021) obtained for the same pictures. Taken together, the findings suggest that that, as intended, moderate to high pleasant or unpleasant affect was induced by nearly all pictures in all included participants.

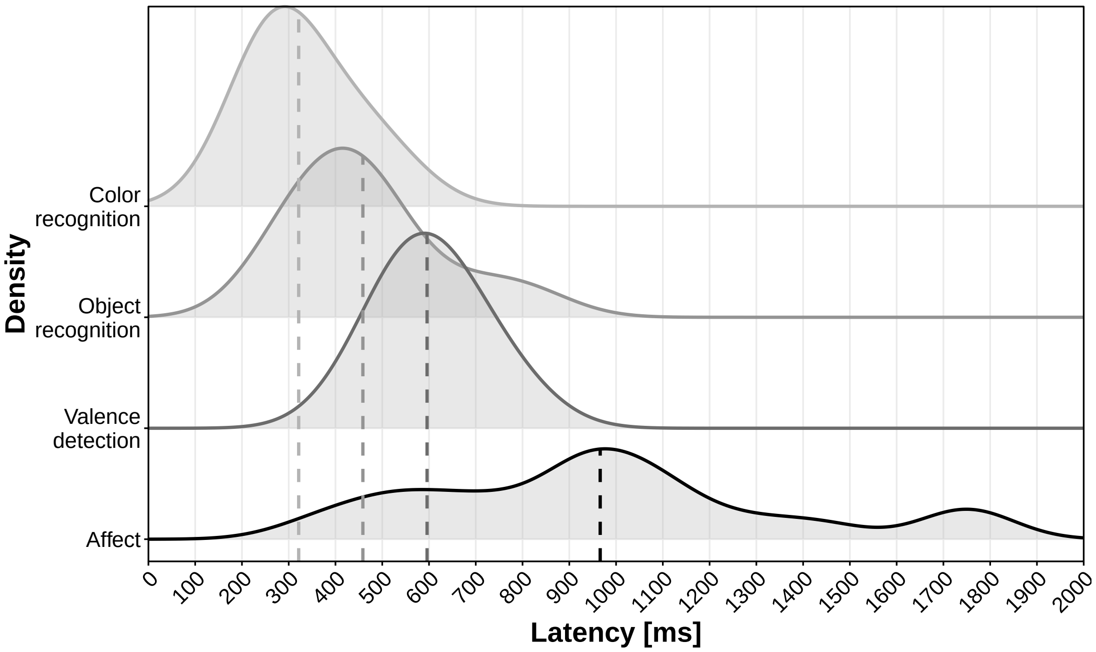
Mean Latencies and Latency Differences

Preliminary ANOVAs conducted on the median latencies revealed that task order, valence order and picture valence had no significant main or interaction effects (see Supplemental Material S1 for details). Therefore, these factors were dropped, meaning that the design was reduced to the factor of main interest, task type. Accordingly, the latency difference hypotheses tested in Experiment 1 (H1, H5, H6, H9) could be tested with paired t-tests. Although the hypotheses are directional, we took a conservative approach and tested them two-tailed.

The average latencies obtained for the four sRT tasks are shown in Figure 1, which also shows the distributions (estimated densities) of the median latencies. As can be seen, the chronological order of the averaged median latencies was color recognition $\prec$ object recognition $\prec$ valence detection $\prec$ affect onset. This is in line with predictions. The average latency of color recognition was 331 ms (excluding one outlying participant with an unreasonably long latency of 3222 ms; 491 ms if she was included), followed by object recognition (465 ms), valence detection (605 ms) and affect (976 ms). Thus, affect was reported 511 ms after object recognition.

Figure 1

Plots of the Densities (bandwidth = 100) of the Medians of the Latencies of Color Recognition, Object Recognition, Valence Detection, and Affect



Note: Vertical dashed lines represent the means of the distributions. One participant was been excluded from the color recognition latencies because of an unreasonably long latency (see text).

Table 1
Between-Subject Correlations for the Medians of Object Recognition, Valence Detection, and Affect (Experiment 1)

	Object recognition	Valence detection	Affect
Object recognition	*.97*		
Valence detection	.38	*.90*	
Affect	.18	.36	*.91*

Note: The italicized coefficients in the diagonal are the reliabilities of the variables, estimated as the intraclass correlation ICC(3, k). See text for more information.

Paired t-tests comparing the adjacent means revealed that all differences were significant. In particular, confirming Hypothesis H1, the latency of object recognition was shorter than the latency at of affect, $t(17) = 5.38$, $p < .001$, $\eta^2_g = .43$ (see also Figure 1). Supporting Hypothesis H5, object recognition was faster than valence detection, $t(17) = 3.93$, $p = .001$, $\eta^2_g = .23$; and supporting Hypothesis H6, valence was detected faster than affect onset, $t(17) = 4.18$, $p = .001$, $\eta^2_g = .31$ (see Figure 3). The median accuracy of the valence judgments was 100%, with 17 of the 18 participants having at least 90% correct.

Supplementary *minF'* analyses (Clark, 1973; Raaijmakers, 1999) revealed that the latency differences remained significant if the items (pictures) were treated as a random rather than a fixed factor (see Supplemental Material S1).

Finally, Hypothesis H9, that color recognition is faster than object recognition, was statistically supported after the outlier in the color recognition latencies was excluded, $t(16) = 2.93$, $p = .010$, $\eta^2_g = .20$ (for details, see Supplemental Material S1).

The latencies of the individual participants corresponded mostly to the mean pattern. The number of participants out of the total 18 whose median RT was shorter on task x than on task y (abbreviated as $x \prec^{N} y$) were: color recognition $\prec^{15}$ object recognition $\prec^{15}$ valence detection $\prec^{16}$ affect onset; and object recognition $\prec^{18}$ affect.

Correlations Among the Latencies

Table 1 shows the Pearson correlations between the median latencies of the three mental events of focal interest. As shown, the correlations between the latencies of object recognition and affect (H2), and between object recognition and valence detection (H7b), were positive as predicted, but low ($r = .18$ and .38) and not significant. The correlation between valence detection and affect ($r = .36$) (H7a) also did not reach significance.

The diagonal of Table 1 contains the intraclass correlations ICC(3, k) of the latencies obtained in the two repetitions of each task (see Method), an index of the reliabilities of the measurements (Shrout & Fleiss, 1979; see also Koo & Li, 2016). ICC(3, k) measures the degree of consistency between the repeated measurements, allowing for systematic mean differences. In the present case, the ICC(3, k) coefficients were essentially identical to ICC(1, k), which measures agreement, indicating that no systematic mean differences between the repeated measurements existed. As can be seen from Table 1, the reliabilities of all latencies were high, ICC(3, k) $\geq$ .90.

Analyses for Individual Pictures

To check whether the findings of main interest depended on the identity of the pictures, the mean differences and correlations between the latencies of object recognition, affect, and valence detection were also calculated for the 24 individual pictures. Descriptively, the predicted pattern of latencies (object recognition $\prec$ valence $\prec$ affect) was found for all 24 pictures. Dependent-sample Yuen t-tests for trimmed means (used to account for RT outliers; see Wilcox, 2016; Yuen, 1974), revealed that the latency differences between object recognition and affect were significant ($p < .05$, two-tailed) in 19 cases, those between object recognition and valence detection in 20 cases, and those between valence detection and affect in 12 cases (for more detail, see Supplemental Material S1).

The median Spearman rank correlation (ρ; used to control for outliers) between the latencies was $\rho = .25$ (range = -.01–.60; 23 of 24 positive) for object recognition and affect, $\rho = .31$ (range = -.17–.73; 22 positive) for object recognition and valence detection; and $\rho = .36$ (range = -.03–.62; 22 positive) for valence detection and affect. For each type of correlation, a quarter (6–7) of the correlations were $> .469$ and thus passed the .05 significance level (two-tailed) for the given sample size.

Discussion

Supporting our predictions, object recognition was faster (511 ms) than affect (H1), valence detection was slower (140 ms) than object recognition (H5) but faster (371 ms) than affect (H6), and color recognition was faster (134 ms) than object recognition (H9). The chronological order object recognition $\prec$ valence detection $\prec$ affect was found for most participants and for all pictures.

The overall between-subject correlations between object recognition, affect, and valence detection were positive as predicted (H2, H7b, H7a), but low and not reliably different from zero. The low size of the correlations cannot be attributed to low reliability of the RT measurements, as these reliabilities were high. Possibly, however, the RTs were contaminated with large systematic errors stemming from interindividual differences in motor or decisional processes. It therefore remains possible that the correlations are indeed positive as predicted, but failed to become significant because of insufficient power. Support for this assumption is provided by the finding that the predicted positive correlations were obtained for nearly all individual pictures, a quarter of which were significant.

In sum, the found latency differences between object recognition, affect, and valence detection support the semantic primacy hypothesis, as well as our hypothesis about the processes underlying valence judgments. The sign of the correlations between the latencies is also in agreement with these hypotheses, but the correlations were low and not reliably different from zero given the sample size.

Experiment 2

Experiment 2 was a replication and extension of Experiment 1. The same latencies were measured, but two experimental manipulations were added: blurring and false-coloring the affective pictures. The rationale for these manipulations and the predictions made for them were described in the introduction.

Method

Participants

The participants were 40 undergraduate psychology students, 34 females and six males (age $M = 21.5$, $SD = 3.2$). Participants were rewarded course credit. One participant was excluded from the data analyses because of low affect ratings (see Results), reducing the final sample size to 39. As in Franikowski et al. (in press), sample size was doubled to increase the chances of detecting the predicted effects of the experimental manipulations.

Stimuli

Stimuli were the pictures used in Experiment 2 of Franikowski et al. (in press). Most of these pictures are identical to those used in the present Experiment 1, but a few unpleasant pictures (e.g., maggots, a decaying carcass) were replaced because they were difficult to recognize when blurred. Pictures were blurred by applying a Gaussian low-pass filter with different blur grades to each picture, and selecting the strongest blur grade at which the object could still be well recognized (Franikowski et al., in press). False-color versions of the sharp and blurred pictures were created by changing the color channels from Red-Green-Blue to Blue-Red-Green. The effect of this color swap is that blue objects or scene parts (e. g, sky, water) appear in a neon-green color, whereas yellow and skin colors appear predominantly violet. The object contours are essentially unaffected. The overall result is that of unnaturally colored but otherwise well-recognizable objects. Junge and Reisenzein (2013, Exp. 2) found that a color manipulation with a somewhat similar visual effect (tinting pictures with neon colors) strongly reduced experienced disgust.

Blurring and false-coloring were implemented as within-subject manipulations. To avoid transfer effects, in particular the facilitation of object recognition, from sharp to blurred pictures, the picture set was divided into two parallel halves (six pleasant and six unpleasant) matched in intensity of evoked affect using the rating data from Franikowski et al. (in press) and Reisenzein and Franikowski (2021), and either the first or second half was blurred for different participants. All pictures were presented in their normal-color and false-color versions. This resulted in 48 affect-evoking stimuli in total.

Procedure

The tasks were kept as much as possible parallel to Experiment 1, but some changes were necessary to accommodate the experimental manipulations. In addition, we added a second practice block, stimulus appearance detection. In this task, the participants were presented with the red color picture from the color recognition task and were instructed to press the “B” key as soon as they detected the onset of the stimulus. Stimulus detection is generally regarded as the simplest RT task (e.g., Jaśkowski, 2014; Woods et al., 2015) and was therefore expected to take even less time than color recognition.

The participants completed the six judgments in the following sequence: Affect ratings, stimulus appearance detection, color recognition, object recognition followed by affective experience or vice versa, and valence detection. The appearance detection and color recognition judgments were made first because they also served to familiarize the participants with the sRT procedure. The valence judgments were made last to provide the participants with repeated (altogether, five) opportunities to determine the valence of each object and store it in memory. All judgments were made in a single session lasting about 45 min.

Design and Data Analyses

The design of Experiment 2 comprised three factors of main interest, task type (object recognition, affect etc.), blurring (sharp or blurred affective pictures), and coloring (normal color or false-color pictures). To control for possible sequence effects, valence order (pleasant or unpleasant pictures judged first) and task order (object recognition followed by affect or vice versa) were counterbalanced, as in Experiment 1. A third control factor in Experiment 2 was blur group (picture set A or picture set B blurred). The participants were randomly assigned to the resulting eight between-conditions, with five subjects in each cell.

The latency difference hypotheses were tested using mixed-design ANOVAs conducted on median RTs. The p-values of predicted effects were uncorrected, whereas those of unpredicted effects were Holm-corrected (Cramer et al., 2016). The correlational hypotheses were tested with multilevel correlation analyses to obtain estimates of both between- and within-subject correlations, and the mediational hypotheses were tested with multilevel path analyses. Mplus 8.1 (Muthén & Muthén, 2017) was used for these analyses.

Results

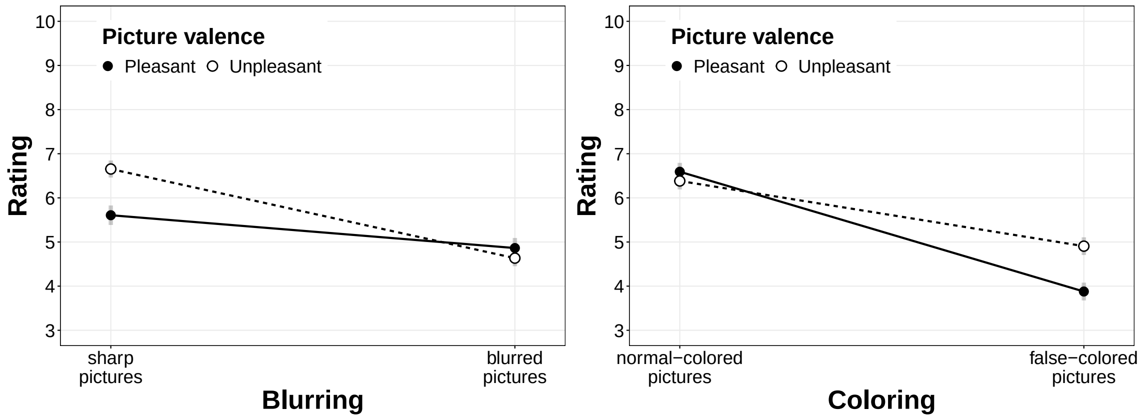
Affect ratings

A 2 (blurring) $\times$ 2 (coloring) $\times$ 2 (picture valence) within-subject $\times$ 2 (blur group: picture set A or B blurred) between-subject ANOVA tested the effects of blurring and coloring on the affect ratings; picture valence was included as a control factor. The ANOVA revealed main effects of blurring, $F(1, 37) = 67.47$, $p < .001$, $\eta^2_g = .13$, and coloring, $F(1, 37) = 360.01$, $p < .001$, $\eta^2_g = .25$, as predicted. In addition, two-way interactions between these factors and picture valence were found, $F(1, 37) = 41.55$, $p < .001$, $\eta^2_g = .03$ for blurring, and $F(1, 37) = 30.14$, $p < .001$, $\eta^2_g = .03$ for coloring. The three-way interaction was also significant, $F(1, 37) = 10.95$, $p = .023$, $\eta^2_g = .01$. No other main or interaction effect was significant, $F \leq 8.22$, $p \geq .068$.

The main effects of blurring and false-coloring replicate findings of Franikowski et al. (in press, Exp. 2) and Junge and Reisenzein (2013, Exp. 2), respectively: Both manipulations reduced the intensity of self-reported affect (see Figure 2). As expected, the effect of false-coloring was stronger ($\eta^2_g = .25$) than the effect of blurring ($\eta^2_g = .13$). The two-way and three-way interactions were due to the fact that the effect of blurring was stronger for unpleasant than pleasant pictures, whereas the effect of false-coloring was stronger for pleasant than unpleasant pictures (see Figure 2). Importantly, the matched picture sets did not differ systematically in their capacity to induce affect, meaning that picture matching was successful. This replicates parallel findings of Franikowski et al. (in press).

Figure 2

Mean Ratings of the Affect-Evoking Pictures



Note: Mean pleasantness and unpleasantness ratings for sharp and blurred pictures (left panel) and for normal-colored and false-colored pictures (right panel). The vertical bars show the 95%-Cousineau-Morey-CIs (Baguley, 2012) of the means.

Overall Latencies and Latency Differences

In agreement with Experiment 1, the rank order of the mean latencies was: color recognition $\prec$ object recognition $\prec$ valence detection $\prec$ affect onset. In addition, the latency of appearance detection (not measured in Experiment 1) was faster than the latency of color recognition. These findings are in agreement with our hypotheses. The size of the latencies for sharp and normal-color pictures (as used in Experiment 1) was also similar to Experiment 1: Appearance detection (273 ms) was fastest, followed by color recognition (295 ms), object recognition (480 ms), valence detection (652 ms) and affect onset (949 ms). Hence affect was reported 469 ms after object recognition. The latencies of the 39 individual participants for the sharp, normal-color pictures largely corresponded to the mean pattern. Using again "$x \prec^{\mathrm{N}} y$" to denote the number of participants for which latency x was smaller than latency y, we found: appearance detection $\prec^{28}$ color recognition $\prec^{36}$ object recognition $\prec^{37}$ valence detection $\prec^{26}$ affect onset; and object recognition $\prec^{38}$ affect. Statistical tests are reported below.

Latencies of Object Recognition versus Affect and the Effects of Picture Blurring and Coloring

Preliminary ANOVAs were performed on the median RTs to explore possible effects of the control variables valence order, task order, picture valence, and blur group. These analyses revealed no significant main effects or interactions involving the control factors with the exception of two-way interactions between picture valence and blurring, and picture valence and coloring (see Supplemental Materials S1). Therefore, picture valence was included into the main analyses, whereas the other control variables were dropped.

As in Experiment 1, we predicted that the latency of object recognition would be shorter than the latency of affect (hypothesis H1). In addition, we predicted that blurring the pictures would delay both object recognition and affect (H3a), whereas false-coloring would only delay affect onset (H4a). These hypotheses were tested simultaneously in a 2 (task: object recognition vs. affect) × 2 (blurring: sharp vs. blurred pictures) × 2 (coloring: normal-color vs. false-color pictures) × 2 (picture valence: pleasant vs. unpleasant pictures) ANOVA conducted on the individual medians of the latencies, which for this purpose were newly aggregated for each participant in the 2^4 cells of this design. Corresponding to the three hypotheses, we predicted main effects for task (H1) and blurring (H3a), and an interaction between task and coloring (H4a). These predicted effects (and the main effect of coloring) were tested with uncorrected p-values whereas the remaining effects were Holm-corrected.

The ANOVA found main effects of task (object recognition vs. affect), $F(1, 38) = 34.41$, $p < .001$, $\eta_g^2 = .17$, blurring, $F(1, 38) = 17.05$, $p < .001$, $\eta_g^2 = .01$, and coloring, $F(1, 38) = 50.44$, $p < .001$, $\eta_g^2 = .03$. In addition, there were interaction effects for task × coloring, $F(1, 38) = 38.09$, $p < .001$, $\eta_g^2 = .02$, coloring × valence, $F(1, 38) = 14.18$, $p = .006$, $\eta_g^2 = .01$, and task × coloring × valence, $F(1, 38) = 18.50$, $p = .001$, $\eta_g^2 = .01$. Figures 3a and 3b show the relevant comparisons side-by-side. Figure 3a shows the means for task × blurring, whereas Figure 3b shows the means for task × coloring × valence.

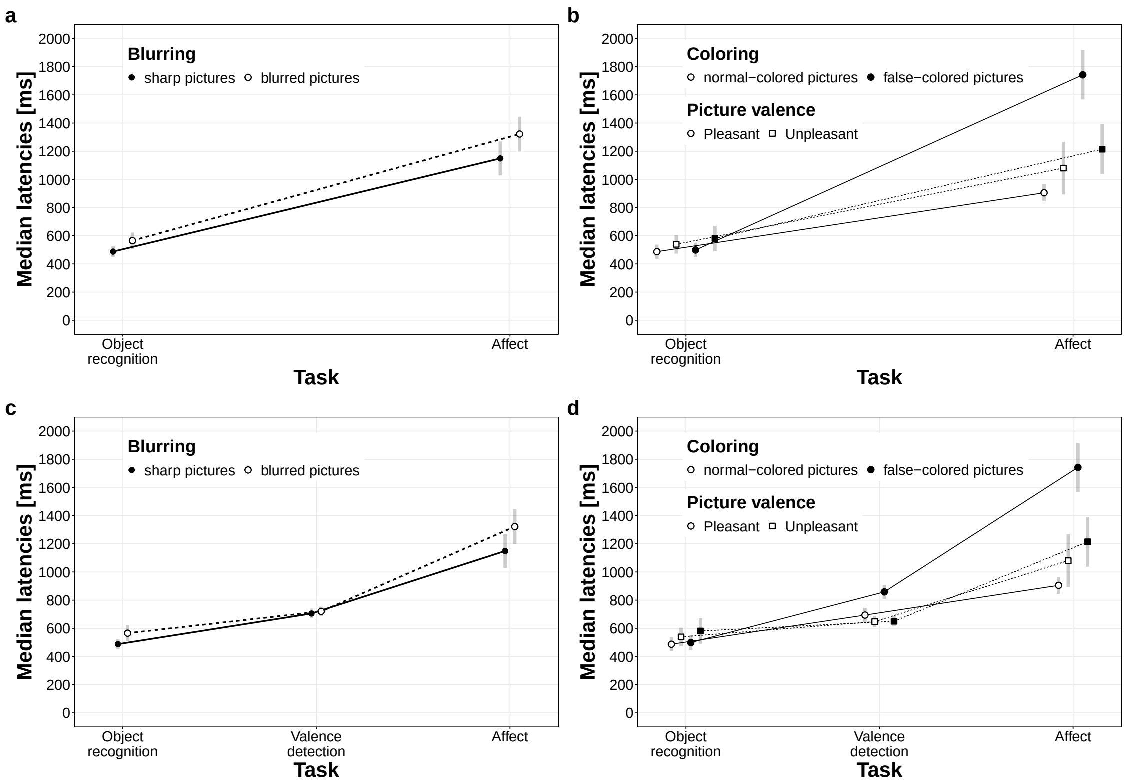
The main effect of task is in line with H1: The latency of affect ($M = 1235$ ms) was significantly longer than the latency of object recognition ($M = 526$ ms). The two interaction effects involving task (task × coloring and task × coloring × valence) reflect the finding that the size of the cognition-affect delay depended on the color and valence of the pictures; however, simple effects analyses using one-factorial ANOVAs confirmed that affect was later than object recognition for each of the four color × valence combinations, $F\text{s} \geq 7.01$, $p\text{s} \leq .012$, $\eta_g^2 \geq .09$. As can be seen from Figure 3b, the cognition-affect lag was smallest for normal-colored pleasant ($M = 418$), followed by normal-colored unpleasant pictures ($M = 541$ ms) and false-colored unpleasant pictures ($M = 633$ ms), and was particularly large for false-colored pleasant pictures ($M = 1243$ ms).

Hypothesis H3a holds that blurring the pictures delays both object recognition and affect. This hypothesis is supported by the main effect of blurring and the lack of interactions involving blurring. Follow-up analyses confirmed that blurring significantly delayed object recognition (M delay = 78 ms), $F(1, 38) = 4.92$, $p = .033$, $\eta_g^2 = .03$; as well as affect (M delay = 173 ms), $F(1, 38) = 13.22$, $p < .001$, $\eta_g^2 = .01$ (see Figure 3a). These findings confirm H3a and replicate corresponding findings of Franikowski et al. (in press).

H4a posits that coloring the pictures delays the latency of perceived affect onset (because it reduces the intensity of experienced affect), but has no or only a small effect on the latency of object recognition. This hypothesis was supported by the significant interaction between the color manipulation and task (object recognition vs. affect; see Figure 3b), and was confirmed by follow-up one-factorial ANOVAs that separately tested the effect of coloring on the latency of object recognition and affect for pleasant and unpleasant pictures. As hypothesized (H4a), false colors delayed the onset of affect for pleasant pictures, $F(1, 38) = 32.57$, $p < .001$, $\eta_g^2 = .214$, M delay = 873 ms, and unpleasant pictures, $F(1, 38) = 4.86$, $p = .034$, $\eta_g^2 = .003$; M delay = 134 ms; but they had only a weak effect on object recognition, that was not

Figure 3

Mean Latencies (Averaged Medians) of the Latencies of Object Recognition, Valence Detection, and Affect



Note: Left column: Mean latencies of object recognition and affect for sharp and blurred pictures without (a) and with (c) valence detection included. Right column: Mean latencies for object recognition and affect for normal-color and false-color pleasant and unpleasant pictures without (b) and with (d) valence detection included. The vertical bars show the 95%-Cousineau-Morey-CIs (Baguley, 2012) of the respective means.

significant for pleasant pictures: $F(1, 38) = 2.37$, $p = .132$, M delay = 12 ms, nor for unpleasant pictures, $F(1, 38) = 2.51$, $p = .121$; M delay = 42 ms (see Figure 3b).

Latency of Valence Detection and the Effects of Picture Blurring and Coloring

To test whether the latency of object recognition is faster than that of valence detection (H5), we computed an ANOVA with the factors task (here: object recognition vs. valence detection) and the control variables blurring, coloring, and picture valence. This ANOVA revealed main effects of task, $F(1, 38) = 22.58$, $p < .001$, $\eta^2_g = .10$, blurring, $F(1, 38) = 6.09$, $p = .018$, $\eta^2_g = .01$, and coloring, $F(1, 38) = 18.96$, $p < .001$, $\eta^2_g = .01$, which were qualified by interactions between task and picture valence, $F(1, 38) = 12.06$, $p = .014$, $\eta^2_g = .03$, and between task, coloring, and valence, $F(1, 38) = 26.71$, $p < .001$, $\eta^2_g = .01$ (see Figures 3c and 3d). No other main or interaction effect was significant, $F \leq 8.03$, $p \geq .073$. Simple-effects analyses revealed that the latency difference between object recognition and valence detection was significant for pleasant pictures, regardless of coloring, Fs ≥ 43.54, ps $\leq .001$, $\eta^2_g \geq .25$ (see Figure 3d). For unpleasant normal-color pictures, the difference was also significant, but weaker, $F(1, 38) = 5.44$, $p = .025$, $\eta^2_g = .067$; and for unpleasant false-color pictures, it was no longer reliable $F(1, 38) = 1.06$, $p = .309$. Overall, then, H5 received reasonable support.

Hypothesis H6, that the latency of affect is longer than the latency of valence detection, was tested using a parallel ANOVA. This ANOVA revealed main effects of task (valence detection vs. affect), $F(1, 38) = 26.61$, $p < .001$, $\eta^2_g = .10$, blurring, $F(1, 38) = 12.41$, $p < .001$, $\eta^2_g = .004$ and coloring, $F(1, 38) = 54.86$, $p < .001$, $\eta^2_g = .03$, as well as two-way interactions between coloring and valence, $F(1, 38) = 20.34$, $p < .001$, $\eta^2_g = .02$, and between task and blurring, $F(1, 38) = 12.22$, $p = .012$, $\eta^2_g = .003$, plus a three-way interaction between task, coloring and picture valence, $F(1, 38) = 11.79$, $p = .013$, $\eta^2_g = .01$ (see again Figures 3c and 3d). No other main or interaction effect was significant, all Fs ≤ 3.12, ps $\geq .681$. The interactions indicate that the size of the lag between valence detection and affect differed for the four combinations of picture valence and coloring. However, simple-effects ANOVAs confirmed that hedonic valence was detected before the onset of affect at all picture valence $\times$ coloring combinations, Fs ≥ 5.95, ps $\leq .020$, $\eta^2_g \geq .06$, as well as in both blurring conditions, Fs ≥ 19.57, ps $< .001$, $\eta^2_g \geq .14$.

Finally, hypothesis H8 posits that blurring and false-coloring both delay the detection of valence. While the results of the two preceding ANOVAs already speak to this hypothesis, we also tested it in a more straightforward manner by conducting a 2 (blurring) $\times$ 2 (coloring) $\times$ 2 (picture valence) ANOVA on the valence judgments. This ANOVA revealed main effects of coloring, $F(1, 38) = 15.47$, $p < .001$, $\eta^2_g = .03$ and picture valence, $F(1, 38) = 13.24$, $p = .001$, $\eta^2_g = .06$, plus an interaction between coloring and picture valence, $F(1, 38) = 25.14$, $p < .001$, $\eta^2_g = .03$; no other effects reached significance. Follow-up t-tests showed that coloring prolonged the latency of valence detection for pleasant pictures, $t(38) = 4.54$, $p < .001$, $\eta^2_g = .08$, but not for unpleasant pictures, $t(38) = 0.24$, $p = .812$. Hence, hypothesis H8 received only very partial support.

Supplementary *minF'* analyses, reported in Supplemental Material S1, revealed that the reported findings were mostly replicated if not only participants but also pictures were treated as random factors in the ANOVAs.

Latencies of Color Recognition and Stimulus Onset Detection

Additional ANOVAs confirmed H9, that the latency of color recognition is significantly shorter than that of object recognition, $F(1, 38) = 48.60$, $p < .001$, $\eta^2_g = .38$, replicating Experiment 1; and H10, that the latency of stimulus appearance detection is significantly shorter than that of color recognition, $F(1, 38) = 8.77$, $p = .005$, $\eta^2_g = .05$ (see Supplemental Material S1 for more detail).

Correlations between the Latencies

Different from Experiment 1, in which no systematic within-subject variation of the latencies was present, the experimental manipulations included in Experiment 2 generated intra-individual variation. This allowed to estimate the between-subject as well as the within-subject correlations between the latencies of object recognition, affect onset, and valence detection (H2 and H7a, H7b). The correlations were estimated using multilevel correlation analysis. To be consistent with the multilevel

Table 2

Correlations between the Latencies of Object Recognition, Valence Detection, and Affect for the Pleasant and Unpleasant Pictures, Experiment 2

Valence	Variable	Object recognition	Valence detection	Affect
Pleasant	Object recognition	*.96*	.47**	.42**
	Valence detection	.24***	*.92*	.36**
	Affect	.14*	.46***	*.89*
Unpleasant	Object recognition	*.90*	.01	-.03
	Valence detection	.02	*.96*	.88***
	Affect	.16***	-.08	*.97*

Note: The italicized coefficients in the diagonal are the between-subject reliabilities of the variables, estimated by the intraclass correlation ICC(3, k). Below the diagonals of the two matrices: within-subject correlations; above the diagonals: between-subject correlations.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

path analysis reported later, the correlations were computed separately for the pleasant and the unpleasant pictures.

To have a greater number of data points for estimating the within-subject correlations, the latencies were aggregated anew by additionally separating the data along the picture repetition factor. This resulted in 8 data points (median latencies) per participant and task, one in each cell of the 2 (coloring) × 2 (blurring) × 2 (picture repetition) design. We specified a full correlation structure among the latencies on both the within- and the between-subjects level, estimated the corresponding covariances using Mplus 8.1 (Muthén & Muthén, 2017) and used the output standardization option to obtain the correlations. The results are shown in Table 2.

A can be seen, hypothesis H2, claiming a positive correlation between the latencies of object recognition and affect, was supported for both picture valences on the within-subject level (for pleasant pictures r = .14; for unpleasant pictures, r = .16), but only for pleasant pictures at the between-subject level (i.e., the subject means, r = .42). The between-subject correlation for unpleasant pictures was close to zero and nonsignificant, r = -.03.

Hypothesis H7a, that the latencies of valence detection and affect onset are positively correlated, was supported on both analysis levels for the pleasant pictures (r_{within} = .46; r_{between} = .36) and on the between-subject level for the unpleasant pictures (r_{between} = .88). The within-subject correlation for unpleasant pictures was not significant and in fact slightly negative (r_{within} = -.08). Hypothesis H7b, that the latencies of valence detection and object recognition are positively correlated, was supported on both analysis levels for the pleasant pictures (r_{within} = .24, r_{between} = .47). In contrast, no association was found for the unpleasant pictures (r_{within} = .02; r_{between} = .01).

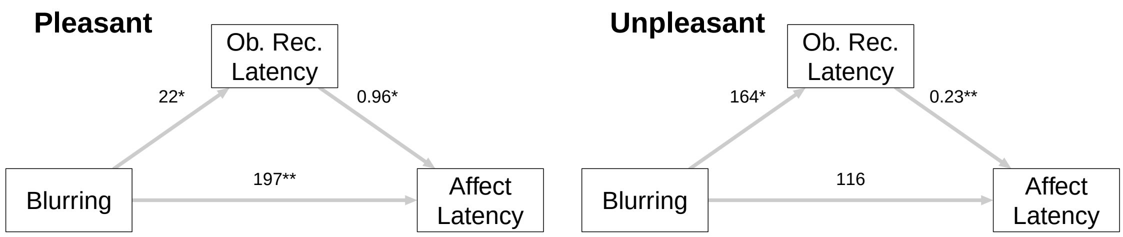
As in Experiment 1, the between-subject reliabilities were high ICC(3, k) $\geq$ .89.

Mediation Analysis of the Blurring Effect

Hypothesis H3b posits that the onset of affect evoked by blurred pictures is delayed because blurred objects need more time to be recognized. To test this hypothesis, we conducted a mediation analysis using the path analytic approach. The estimated causal model is shown in Figure 4. The model was separately estimated for pleasant and unpleasant pictures to account for possible moderating effects of picture valence. The paths were estimated with multilevel structural equation modeling (MSEM; Preacher et al., 2010), using Mplus 8.1 (Muthén & Muthén, 2017). In MSEM, the predictor variables are decomposed during estimation into two latent variable parts, a within- and a between part (Muthén & Muthén, 2017); however, because blurring was a pure within-variable, it was specified as such in Mplus. The same data as in the multilevel correlation analysis (with 8 data points per subject) were entered into the analyses to improve the estimation of the within-paths. Preliminary analyses revealed that the random-slope coefficients had very low variance, indicating that none of the paths required a random slope; therefore, only random intercepts were estimated. On the between-level (the relations between the means of the variables), the path from the latency of object recognition to the latency of affect was also estimated; this path corresponds to the between-subject correlation between these variables, reported above. Inclusion of this path is necessary to obtain a clear separation of the within- and between subject components of the path from the latency of object recognition to the latency of affect (Hoffman, 2019).

Figure 4

Path Model for the Effects of Picture Blurring, Experiment 2



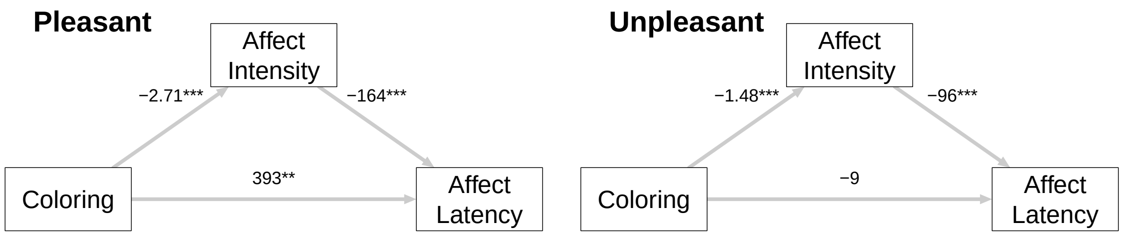
Note: Numbers next to the arrows are unstandardized path coefficients together with their significance levels ($^{*}p < .05$, $^{**}p < .01$). Separate mediation analyses were conducted for the pleasant (left diagram) and unpleasant (right diagram) pictures. Ob. Rec. = object recognition.

The path analysis revealed a significant path from blurring to the latency of object recognition for both pleasant pictures, β = 22, 95% CI = [4, 40], $p = .018$ (according to the critical ratio z-test) and unpleasant pictures, β = 164, 95% CI = [31, 297], $p = .016$, meaning that the participants needed 22 ms (pleasant) and 164 ms (unpleasant) more to recognize the depicted objects if the pictures were blurred (see Figure 4). The path from the latency of object recognition to affect was also significant for pleasant pictures, β = 0.96, 95% CI = [0.18, 1.75], $p = .016$, and unpleasant pictures, β = 0.23, 95% CI = [0.08, 0.38], $p = .002$. These findings are consistent with partial mediation (Woody, 2011). In addition, the direct path from blurring to the latency of affect was significant for pleasant pictures, β = 197, 95% CI = [55, 338], $p = .006$, but not for unpleasant pictures β = 116, 95% CI = [-58, 289], $p = .191$. We can therefore conclude that hypothesis H3b, which claims that the effect of blurring on affect latency is partly mediated by delayed object recognition, was supported for both pleasant and unpleasant pictures.

Mediation Analysis of the False-Color Effect

According to hypothesis H4b, coloring the pictures delayed affect onset because it reduced the intensity of experienced affect, and weaker internal signals need more time to reach the decision criterion (cf. Miller & Schwarz, 2006). This mediation hypothesis was tested in a second multilevel path analysis, again conducted separately for pleasant and unpleasant pictures to take account of the interactions of task with picture valence. Coloring was modeled as a pure within-variable. The estimated causal model is shown in Figure 5. Preliminary analyses revealed that a random-intercept model was acceptable for the unpleasant pictures, whereas for the pleasant pictures, the effects of color required random slopes. The mediator variable, affect intensity, consisted of the mean pleasantness ratings of each participant of the sharp and blurred, normal and false-colored pictures, collected at the beginning of the experiment, on the assumption that these affect intensities would be proportional to those evoked by the pictures in the sRT tasks. On the between level, the path from affect intensity to affect latency was included (see Hoffman, 2019).

Figure 5
Path Model for the Effects of Picture Coloring, Experiment 2



Note: Numbers next to the arrows are unstandardized path coefficients together with their significance levels ($^{**}p < .01$, $^{***}p < .001$). Separate mediation models were estimated for the pleasant (left diagram) and unpleasant (right diagram) pictures.

This path analysis revealed a significant path from coloring to affect intensity for both pleasant pictures, β = -2.71, 95% CI = [-3.09, -2.34], $p < .001$, and unpleasant pictures, β = -1.48, 95% CI = [-1.73, -1.22], $p < .001$. The path from affect intensity to affect latency was also significant for pleasant pictures, β = -164, 95% CI = [-243, -85], $p < .001$, and unpleasant pictures, β = -96, 95% CI = [-141, -51], $p < .001$ (see Figure 5). False colors reduced the intensity of affect evoked by both pleasant and unpleasant pictures, and higher affect intensity reduced (or lower intensity increased) the latency of affect. In addition, for pleasant pictures, there was a significant direct path from coloring to affect latency, β = 393, 95% CI = [151, 635], $p = .001$, suggesting that false-color pictures also delayed affect via another route. However, for unpleasant pictures, the direct path was not significant, β = -9, 95% CI = [-146, 129], $p = .899$. Overall, then, hypothesis H4b, according to which false-color objects delay affect onset because they reduce the intensity of experienced affect, was supported, although the mediation for pleasant pictures was only partial.

Analyses for Individual Pictures

To check whether the findings of the analyses for individual pictures conducted in Experiment 1 can be replicated, we repeated these analyses for the sharp, normal-colored pictures. The results largely replicated the findings obtained in Experiment 1. Details are reported in Supplemental Material S1.

Discussion

The chronological order of the latencies measured in Experiment 1 was replicated in Experiment 2: color recognition $\prec$ object recognition $\prec$ valence detection $\prec$ affect. The average latencies of these judgments for sharp, normal-color pictures (corresponding to those used in Experiment 1) were also similar to Experiment 1. In addition, stimulus detection was found to be faster than color recognition, as predicted. All mean differences were statistically significant, and the pairwise differences were for the most part also found on the level of individual participants. Furthermore, the latency difference between the focal judgments object recognition, affect onset and valence detection, were found for most individual pictures.

Multilevel correlation analyses provided partial support for hypotheses H2 and H7. A positive correlation between the latencies of object recognition and affect (H2) was found for both picture valences on the within-subject level and for pleasant pictures at the between-subject level. The predicted positive correlation between the latencies of valence detection and affect onset (H7a) was found on both analysis levels for the pleasant pictures, and on the between-subject level for unpleasant pictures. The predicted positive correlation between the latencies of valence detection and object recognition (H7b) was supported on both analysis levels for the pleasant pictures, but not for unpleasant pictures. H8 received only very partial support: False-coloring prolonged the latency of valence detection for pleasant but not unpleasant pictures, and blurring had no effect.

Mediation analyses provided support for hypothesis H3b, that the effect of blurring on the latency of affect is at least partly mediated by delayed object recognition; as well as for hypothesis H4b, that false-color pictures delay affect because they reduce the intensity of affect. However, for pleasant pictures, coloring delayed affect also by a different route.

In sum, the findings of Experiment 2 further support the semantic primacy hypothesis. In addition, they support the hypothesis that the latency of affect is also influenced by affect intensity, and they shed some light on the processes underlying judgments of hedonic valence.

General Discussion

Two experiments used sRT tasks to estimate and compare the latencies of, in particular, object recognition, affect, and valence detection. We now discuss the findings as they pertain to the three research goals described in the introduction.

Test of the Semantic Primacy Hypothesis

The first and main goal of the experiments was to re-test the semantic primacy hypothesis in the sRT paradigm. Three hypotheses derived from semantic primacy were tested, all of which received at least partial support.

First, the RT latency of object recognition was significantly shorter than the latency of object recognition (both experiments). Descriptively, the predicted cognition-affect delay was found for all participants of Experiment 1 and for most of Experiment 2, and for each of the 24 individual pictures in both experiments. Second, the latencies of object recognition and affect correlated positively on both the within-subject (Exp. 2) and the between-subject level (Exp. 1 and 2), and these correlations were significant on the within-level for pleasant and unpleasant pictures, as well as on the between-level for pleasant pictures (Exp. 2). Third, picture blurring delayed object recognition and affect, and the effect of blurring on affect latency was partly mediated by delayed object recognition (Exp. 2).

The found cognition-affect lag replicates results obtained by Reisenzein and Franikowski (2021) with temporal order and simultaneity judgments, and by Franikowski et al. (in press) with the rotating spot method. It also replicates the findings of previous chronometric tests of the semantic primacy hypothesis using sRT (Nakashima, 1909a, 1909b; Nummenmaa et al., 2010 using valence judgments). The absolute latencies of object recognition and affect are in fact close to those found by Nakashima (1909a; 1909b) in his sRT experiments.

The positive correlations between the latencies of object recognition and affect replicate parallel between-subject correlations obtained by Reisenzein and Franikowski (2021) with temporal order and simultaneity judgments and by Franikowski et al. (in press) with the rotating spot method, as well as by Nummenmaa et al. (2010; Exp. 3) with RT measurements in a by-picture analysis.

Finally, the delaying effect of blurring on the onset of affect and the partial mediation of this effect by delayed object recognition replicates findings by Franikowski et al. (in press) in the rotating spot paradigm.

In sum, both the present and previous reaction time studies and temporal judgment studies using different latency measurement methods (TOJ, SJ, and the rotating spot method), consistently support the *temporal primacy* of object recognition over affect. As mentioned in the introduction, this conclusion is consistent with the longer latency of the affective modulation of the LPP (see e.g., Hajcak & Foti, 2020) compared to the latency of object recognition (see e.g., Fabre-Thorpe, 2011). Because causes must precede their effects, these findings also support (in the sense of "are consistent with") the *causal primacy* of object recognition over affect. Additional support for causal primacy is the positive correlation between the latencies of object recognition and affect. However, the strongest evidence for the causal primacy of cognition over affect in the latency measurement paradigm are the findings that the experimental delay of object recognition by picture blurring also delays the onset of affect, and that the blurring→affect latency effect is, at least in part, mediated by delayed object recognition. These findings have now been replicated across two different latency measurement methods (sRT tasks and the rotating spot method). We therefore feel reasonably confident that these effects are robust. In addition, the delaying effect of picture blurring on the latency of affect is consistent with the findings of a recent ERP study by Codispoti et al. (2021; see also, Schupp et al., 2008) that the affective modulation of the LPP is delayed if affective pictures are presented in a degraded format that makes object recognition more difficult.

There is, however, one noteworthy difference between the present findings and those of the studies using temporal judgments (Franikowski et al., in press; Reisenzein & Franikowski, 2021): The delays between object recognition and affect onset (on average, 511 ms in Experiment 1 and 469 in Experiment 2, for sharp, normal-color pictures) were about four times larger than in the TJ experiments (120–150 ms). Because we used the same (Experiment 1) or largely the same (Experiment 2) stimuli and similar presentation conditions as in the previous experiments and tested participants from the same subject pool, these differences can be attributed to the latency measurement methods (TJ vs. sRT tasks).

As it turns out, this between-method dissociation is consistent with previous research which found that the time lags between the onsets of different sensations (e.g., between detecting loud and soft tones) are 2–4 times larger if the latencies are measured with sRTs than with TJs (see Cardoso-Leite & Gorea, 2010; Jaśkowski, 2014). A parsimonious explanation of this difference, proposed by several authors, is that participants use a higher criterion for deciding on the presence of a signal in sRT tasks than in TJ tasks (Cardoso-Leite & Gorea, 2010; Miller & Schwarz, 2006). This means, for example, that a higher intensity of pleasure is required to judge that pleasure is present in the RT task, than in TJ tasks. If one assumes that signal intensity increases approximately linearly with time, a higher criterion will result in longer latencies. Miller and Schwarz (2006) demonstrated via simulations that a high decision criterion is optimal in sRT tasks, because it allows to simultaneously maximize speed and minimize false alarms. In contrast, a lower criterion is optimal in TOJ tasks, where speed plays no role, because it allows to maximize precision.

The cognition-affect lag of about 500 ms found in our sRT experiments was also much larger than that found in the sRT studies by Nummenmaa et al. (2010; 30–80 ms). Our findings suggest that the main reason for this difference is that Nummenmaa et al. (2010) measured the latency of valence detection rather than affect onset. For valence judgments, the cognition-affect lag in our studies was much closer to the findings of Nummenmaa et al. (2010): 168 ms in Exp. 1 and 172 ms in Exp. 2. The remaining latency difference can be explained by procedural differences. Probably most important, in the experiments by Nummenmaa et al. (2010) the stimuli were presented for 30 ms only, whereas in our experiments, they remained on the screen for at least 2000 ms. The longer presentation times allow participants to delay responding until they are certain about the valence of the stimulus. In contrast, if the stimulus is only presented for 30 ms, certainty about its valence is probably often not reached, but participants respond immediately nonetheless because no additional information about the stimulus is forthcoming.

Affect Intensity as a Determinant of Affect Latency

The second goal of our experiments was to test the hypothesis that the latency of affect onset also depends on the intensity of affect. Supporting the hypothesis, a manipulation that reduced picture-evoked affect but not object recognition, false-coloring the pictures, also delayed affect onset, and this effect was found to be partly (pleasant pictures) or completely (unpleasant pictures) mediated by reduced affect intensity.

Although these findings are consistent with our predictions, some caution is warranted. First, the findings need to be replicated. Second, the significant and substantial direct effect of coloring on affect latency found for pleasant pictures indicates that coloring influenced affect latency not only by reducing affect intensity. One possibility is that the false-color pictures also changed the quality of the affect evoked by the pictures. For example, the false-color pictures could have evoked mixed feelings (see e.g., Schimmack, 2005). To illustrate, a laughing child with a bluish face may evoke a pleasant feeling caused by the nature of the object (laughing child), but also a negative feeling because of the child's unnatural skin color. In this case, the participants may need extra time to focus attention on, and respond to, the pleasant feeling (as they were instructed to do). Another possibility is that the affect-generating mechanisms need more time to compute affect for false-colored objects, because the category information (e.g., child) and information about unusual object features (e.g., bluish face) is difficult to integrate. A cleaner manipulation of affect intensity that leaves object recognition untouched, may be the manipulation of picture size (De Cesarei & Codispoti, 2008; Junge & Reisenzein, 2013).

Processes Underlying Valence Judgments

The third goal of our experiments was to test the hypothesis that valence judgments of affective pictures are partly based on picture-evoked feelings, and partly on the retrieval of valence information stored in object schemas. Some support for this hypothesis was obtained. Specifically, in agreement with the hypothesis, the average latency of valence detection was in between the latencies of object recognition and affect onset. This finding also implies that Nummenmaa et al.'s (2010) findings (object recognition $\prec$ valence detection) were replicated for a different set of stimuli, a much longer stimulus presentation time, and a different object recognition question. In addition, positive correlations between these latencies and the latency of object recognition were found for pleasant pictures, as well as for unpleasant pictures on the between-subject level (Exp. 2). Finally, the false-color manipulation delayed valence detection for pleasant but not unpleasant pictures, while blurring had no significant effect.

In interpreting these findings, it should be kept in mind that, according to the proposed model of the processes underlying valence judgments, the size of the predicted effects of coloring and blurring on the latency of valence detection depends on the proportions of feeling-based vs. memory-based valence judgments. If the valence judgments of a set of pictures are mostly based on stored valence information, then variables that influence the intensity of picture-evoked affect can only have small effects. The finding that the average latency of affect was much longer than the latency of the valence judgments could be interpreted as an indication that the valence judgments – which, to recall, were presented at the end of the experiment – were for the greater part memory-based.

Again, however, caution is indicated. First, the valence judgment task was a choice RT task, rather than a simple RT task; and second, it was presented at the end of the experiment. Therefore, the shorter latency of valence detection compared to affect could have been due to differences between simple and choice RT tasks, or to practice effects. However, other factors constant, response latencies in choice RT tasks are typically longer than those in simple RT tasks (e.g., Miller & Low, 2001).

As regards the practice hypothesis, our model of the processes underlying valence judgments actually predicts a specific kind of practice effect: Valence judgments become faster with repeated presentations of the same stimuli because the participants can increasingly rely on stored valence information. A third alternative explanation is that the participants in the 1AFC task did base their responses on the picture-evoked feelings, but used a lower decision criterion, analogous to what has been proposed for TOJ tasks (Miller & Schwarz, 2006). A test of this hypothesis must be left to future research.

To conclude, the present studies further support the proposal that the methods of latency measurement of subjective experiences, originally developed in introspective psychology (e.g., Wundt, 1897) and refined in modern cognitive psychology (e.g., Neumann & Niepel, 2004; Posner, 1986; Schneider & Bavelier, 1973) are also useful in emotion psychology.

References

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods, 44*(1), 158–175. https://doi.org/10.3758/s13428-011-0123-7

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379–384. https://doi.org/10.3758/BF03192707

Barrett, L. F. (1996). Hedonic tone, perceived arousal, and item desirability: Three components of self-reported mood. *Cognition & Emotion, 10*(1), 47–68. https://doi.org/10.1080/026999396380385

Bradley, M. M. (2009). *Natural selective attention: Orienting and emotion. Psychophysiology, 46*(1), 1–11.

Bramão, I., Reis, A., Petersson, K. M., & Faísca, L. (2011). The role of color information on object recognition: A review and meta-analysis. *Acta Psychologica, 138*(1), 244–253. https://doi.org/10.1016/j.actpsy.2011.06.010

Cardoso-Leite, P., & Gorea, A. (2010). On the perceptual/motor dissociation: A review of concepts, theory, experimental paradigms and data interpretations. *Seeing and Perceiving, 23*(2), 89–151. https://doi.org/10.1163/187847510x503588

Clark, H. H. (1973). The language–as–fixed–effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*(4), 335–359. https://doi.org/10.1016/S0022-5371(73)80014-3

Clore, G. L., & Byrne, D. (1974). A reinforcement-affect model of attraction. In T. L. Huston (Ed.), *Foundations of Interpersonal Attraction* (pp. 143–170). Academic Press.

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology, 46*(4), 731–738. https://doi.org/10.1111/j.1469-8986.2009.00804.x

Codispoti, M., Ferrari, V., & Bradley, M. M. (2006). Repetitive picture processing: Autonomic and cortical correlates. *Brain Research, 1068*(1), 213–220. https://doi.org/10.1016/j.brainres.2005.11.009

Codispoti, M., Micucci, A., & De Cesarei, A. (2021). Time will tell: Object categorization and emotional engagement during processing of degraded natural scenes. *Psychophysiology, 58*(1), e13704. https://doi.org/10.1111/psyp.13704

Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., Waldorp, L. J., & Wagenmakers, E.-J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review, 23*(2), 640–647. https://doi.org/10.3758/s13423-015-0913-5

Crane, T. (1992). The nonconceptual content of experience. In T. Crane (Ed.), *The contents of experience: Essays on perception* (pp. 136–157). Cambridge University Press.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology, 52*(2), 95–111. https://doi.org/10.1016/s0301-0511(99)00044-7

De Cesarei, A., & Codispoti, M. (2008). Fuzzy picture processing: Effects of size reduction and blurring on emotional processing. *Emotion, 8*(3), 352. https://doi.org/10.1037/1528-3542.8.3.352

De Cesarei, A., & Codispoti, M. (2011). Scene identification and emotional response: Which spatial frequencies are critical? *Journal of Neuroscience, 31*(47), 17052–17057. https://doi.org/10.1523/JNEUROSCI.3745-11.2011

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86–89.

Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology, 2*, 243. https://doi.org/10.3389/fpsyg.2011.00243

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In J. Krosnick & R. E. Petty (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Lawrence Erlbaum Associates.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition, 25*(5), 603–637. https://doi.org/10.1521%2Fsoco.2007.25.5.603

Ferrari, V., Mastria, S., & Codispoti, M. (2020). The interplay between attention and long-term memory in affective habituation. *Psychophysiology, 57*(6), e13572. https://doi.org/10.1111/psyp.13572

Fishbein, M. (1963). An investigation of the relationships between beliefs about an object and the attitude toward that object. *Human Relations, 16*(3), 233–239. https://doi.org/10.1177/001872676301600302

Fiske, S. T. (1982). Schema-triggered affect: Applications to social perception. In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition: The seventeenth annual Carnegie Mellon Symposium on Cognition* (pp. 55– 76). Erlbaum.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117*(1), 39–66. https://doi.org/10.1037/0033-2909.117.1.39

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116–124.

Franikowski, P., Kriegeskorte, L.-S., & Reisenzein, R. (2021). Perceptual latencies of object recognition and affect measured with the rotating spot method: Chronometric evidence for semantic primacy. *Emotion, 21*(8), 1744–1759. https://doi.org/10.1037/emo0000991

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology, 57*(7), e13570. https://doi.org/10.1111/psyp.13570

Hedger, N., Adams, W. J., & Garner, M. (2015). Fearful faces have a sensory advantage in the competition for awareness. *Journal of Experimental Psychology: Human Perception and Performance, 41*(6), 1748–1757. https://doi.org/10.1037/xhp0000127

Hoffman, L. (2019). On the interpretation of parameters in multivariate multilevel models across different combinations of model specification and estimation. *Advances in Methods and Practices in Psychological Science, 2*(3), 288–311. https://doi.org/10.1177/2515245919842770

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70. http://www.jstor.org/stable/4615733

Itkes, O., Kimchi, R., Haj-Ali, H., Shapiro, A., & Kron, A. (2017). Dissociating affective and semantic valence. *Journal of Experimental Psychology: General, 146*(7), 924–942. https://doi.org/10.1037/xge0000291

James, W. (1890). *The Principles of Psychology.* Holt.

Jaśkowski, P. (2014). What determines simultaneity and order perception? In V. Arstila & D. Lloyd (Eds.), *Subjective time: The philosophy, psychology, and neuroscience of temporality* (pp. 379–407). MIT Press.

Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision, 3*(7), 499–512. https://doi.org/10.1167/3.7.4

Junge, M., & Reisenzein, R. (2013). Indirect scaling methods for testing quantitative emotion theories. *Cognition & Emotion, 27*(7), 1247–1275. https://doi.org/10.1080/02699931.2013.782267

Kaplan, M. F., & Anderson, N. H. (1973). Information integration theory and reinforcement theory as approaches to interpersonal attraction. *Journal of Personality and Social Psychology, 28*(3), 301–312. https://doi.org/10.1037/h0035112

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior Research Methods, 49*(2), 457–470.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention, 1*, 39–58.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261–273.

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist, 37*(9), 1019–1024. https://doi.org/10.1037/0003-066X.37.9.1019

Lazarus, R. S. (1995). Vexing research problems inherent in cognitive-mediational theories of emotion – and some solutions. *Psychological Inquiry, 6*, 183–196.

Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General, 144*(2), 339–365. https://doi.org/10.1037/xge0000040

LeDoux, J. E. (1998). *The emotional brain: The mysterious underpinning of emotional life.* Simon & Schuster.

Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavioral disorders* (Vol. 1, pp. 333–378). Ronald Press.

Leys, R. (2017). *The ascent of affect: Genealogy and critique.* University of Chicago Press.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain, 106*(3), 623–642. https://doi.org/10.1093/brain/106.3.623

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*(3), 203–229. https://doi.org/10.1037/a0012869

Mackie, J. L. (1974). *The cement of the universe: A study of causation.* Clarendon Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* Freeman.

Meinong, A. (1894). *Psychologisch-ethische Untersuchungen zur Werttheorie. [Psychological-ethical investigations in value theory].* Leuschner & Lubensky.

Micucci, A., Ferrari, V., De Cesarei, A., & Codispoti, M. (2020). Contextual modulation of emotional distraction: Attentional capture and motivational significance. *Journal of Cognitive Neuroscience, 32*(4), 621–633. https://doi.org/10.1162/jocn_a_01505

Miller, J. O., & Low, K. (2001). Motor processes in simple, go/no-go, and choice reaction time tasks: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance, 27*(2), 266–289. https://doi.org/10.1037/0096-1523.27.2.266

Miller, J. O., & Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: A diffusion model approach. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 394. https://doi.org/10.1037/0096-1523.32.2.394

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393*(6684), 467–470. https://doi.org/10.1038/30976

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review, 8*(3), 454–469. https://doi.org/10.3758/bf03196180

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723–739.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén &

Muthén.

Myrick, J. G. (2015). Emotion regulation, procrastination, and watching cat videos online: Who watches Internet cats, why, and to what effect? *Computers in Human Behavior, 52*, 168–176. https://doi.org/10.1016/j.chb.2015.06.001

Nakashima, T. (1909a). Contributions to the study of the affective processes. *American Journal of Psychology, 20*(2), 157–193.

Nakashima, T. (1909b). Time-relations of the affective processes. *Psychological Review, 16*(5), 303–339. https://doi.org/10.1037/h0075294

Neumann, O., & Niepel, M. (2004). Timing of perception and perception of time. In C. Kaernbach, E. Schröger, & E. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 245–269). Lawrence Erlbaum.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General, 139*(2), 222–246. https://doi.org/10.1037/a0018858

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41*(2), 176–210. https://doi.org/10.1006/cogp.1999.0728

Ooms, J. (2020). *Magick: Advanced Graphics and Image-Processing in R* (Version 2.3) [Computer software].

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* Cambridge University Press.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Towards an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522. https://psycnet.apa.org/doi/10.1037/0033-295X.108.3.483

Or, C. C.-F., Retter, T. L., & Rossion, B. (2019). The contribution of color information to rapid face categorization in natural scenes. *Journal of Vision, 19*(5), 20–20. https://doi.org/10.1167/19.5.20

Osgood, C. h. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* University of Illinois Press.

Peira, N., Golkar, A., Öhman, A., Anders, S., & Wiens, S. (2012). Emotional responses in spider fear are closely related to picture awareness. *Cognition & Emotion, 26*(2), 252–260. https://doi.org/10.1080/02699931.2011.579087

Pessoa, L., Japee, S., & Ungerleider, L. G. (2005). Visual awareness and the detection of fearful faces. *Emotion, 5*(2), 243–247. https://doi.org/10.1037/1528-3542.5.2.243

Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition, 16*, 241–254. https://doi.org/10.1016/j.concog.2006.09.002

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. https://doi.org/10.1037/a0020141

R Core Team. (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Raaijmakers, J. G., Schrijnemakers, J. M., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41*(3), 416–426. https://doi.org/10.1006/jmla.1999.2650

Reisenzein, R. (2009). Emotions as metarepresentational states of mind: Naturalizing the belief–desire theory of emotion. *Cognitive Systems Research, 10*(1), 6–20. https://doi.org/10.1016/j.cogsys.2008.03.001

Reisenzein, R. (2012). What is an emotion in the Belief–Desire Theory of emotion? In F. Paglieri, L. Tummolini, R. Falcone, & M. Miceli (Eds.), *The goals of cognition. Essays in honor of Cristiano Castelfranchi* (pp. 181–211). College Publications.

Reisenzein, R., & Franikowski, P. (2021). *On the latency of object recognition and affect: Evidence from temporal order and simultaneity judgments* [Manuscript currently under review].

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172.

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3459–3474. https://doi.org/10.1098/rstb.2009.0141

Schimmack, U. (2005). Response latencies of pleasure and displeasure ratings: Further evidence for mixed feelings. *Cognition & Emotion, 19*(5), 671–691. https://doi.org/10.1080/02699930541000020

Schupp, H. T., Stockburger, J., Schmälzle, R., Bublatzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: Brain potentials and stimulus identification. *NeuroReport, 19*(2), 167–171. https://doi.org/10.1097/wnr.0b013e3282f4aa42

Schwarz, N., & Clore, G. L. (2007). Feelings and phenomenal experiences. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (p. 385–407). The Guilford Press.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420–428. https://psycnet.apa.org/doi/10.1037/0033-2909.86.2.420

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *Afex: Analysis of Factorial Experiments* (Version 0.27-2) [Computer software].

Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic processing precedes affect retrieval: The neurological case for cognitive primacy in visual processing. *Review of General Psychology, 10*(1), 41–55. https://doi.org/10.1037/1089-2680.10.1.41

Weiß, K., & Scharlau, I. (2010). Temporal order judgments: A sensitive measure for measuring perceptual latency? In A. Bastianelli & G. Vidotto (Eds.), *Fechner Day 2010: Proceedings of the 26th Annual Meeting of the International Society for Psychophysics* (pp. 467–472). The International Society for Psychophysics.

Wells, F. L. (1925). Reactions to visual stimuli in affective settings. *Journal of Experimental Psychology, 8*, 64–76. https://doi.org/10.1037/h0066415

Wells, F. L. (1929). Reaction-times to affects accompanying smell stimuli. *The American Journal of Psychology, 41*(1), 83–86.

Whelan, R. (2008). Effective Analysis of Reaction Time Data. *The Psychological Record, 58*(3), 475–482. https://doi.org/10.1007/bf03395630

Wilcox, R. R. (2016). *Introduction to robust estimation and hypothesis testing* (4th edition). Elsevier.

Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., Reed, B. (2015). Factors influencing the latency of simple reaction time. *Frontiers in Human Neuroscience, 9*. https://doi.org/10.3389/fnhum.2015.00131

Woody, E. (2011). An SEM perspective on evaluating mediation: What every clinical researcher needs to know. *Journal of Experimental Psychopathology, 2*(2), 210–251. https://doi.org/10.5127/jep.010410

World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. (2013). *JAMA, 310*(20), 2191. https://doi.org/10.1001/jama.2013.281053

Yuen, K. K. (1974). The two-sample trimmed t for unequal population variances. *Biometrika, 61*(1), 165–170. https://doi.org/10.2307/2334299

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Corrigendum to Publication C

Please note that this is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Aside from minor changes, we introduced some more detail and improved the writing during the copy editing process. These adaptations must not be introduced in the document due to copyright conflicts. However, major changes in comparison to this version of the document will be summarized here.

- Section **Discussion** of Experiment 1, paragraph 2: “The overall between-subject correlations between object recognition, affect, and valence detection were positive as predicted (H2, H7b, H7a) [...]” was changed to: “The between-subjects correlations between the overall (i.e., across-pictures) latencies of object recognition, affect, and valence detection were positive as predicted (H2, H7b, and H7a) [...]”
- Section **Mediation Analysis of the Blurring Effect** of the Results section for Experiment 2, paragraph 2: first entry of “β” was extended to: “(unstandardized) β” (all coefficients were unstandardized)
- Section **Test of the Semantic Primacy Hypothesis** of the General Discussion, paragraph 2: “First, the RT latency of object recognition was significantly shorter than the latency of object recognition (both experiments).” was corrected to: “First, the sRT latency of object recognition was significantly shorter than the latency of affect (both experiments).”; moreover, “24 individual pictures” was corrected to: “24 unmodified pictures”

The final article is available, upon publication, at: https://doi.org/10.1037/emo0001092.supp

On the Latency of Object Recognition and Affect: Evidence from Speeded Reaction Time Tasks

Rainer Reisenzein and Philipp Franikowski

Supplemental Materials S1: Supplementary Analyses

Experiment 1

Preliminary ANOVA

Two ANOVAs were computed. The first, which comprised all tasks, was a 4 (task: color recognition, object recognition, valence detection, or affect) × 2 (task order) × 2 (valence order) ANOVA. The second, which comprised only the tasks for the affective pictures, was a 3 (task: object recognition, valence detection, or affect) × 2 (valence: pleasant vs. unpleasant pictures) × 2 (task order) × 2 (valence order) ANOVA. Task and valence were within-subject factors, whereas task order and valence order were between-subject factors. The p-values of the main and interaction effects involving the control variables were Holm-corrected (Holm, 1979) following the recommendation of Cramer et al. (2016) to protect against the inflation of the significance level in multiway ANOVAs. The main effect of task was uncorrected, because it corresponded to the a priori hypothesis. Both ANOVAs revealed a main effect of task: First ANOVA: $F(1.87, 26.11) = 5.64$, $p = .010$, $\eta_g^2 = .21$; second ANOVA: $F(1.20, 16.86) = 21.75$, $p < .001$, $\eta_g^2 = .35$. All other main and interaction effects did not reach significance: first ANOVA: $F \leq 2.29$, $p \geq .614$; second ANOVA: $F \leq 3.47$, $p \geq .942$.

Latencies of Color Recognition

A t-test comparing the latencies of color recognition and object recognition revealed no significant difference, $t(17) = -0.16$, $p = .872$. Inspection of the data indicated, however, that this finding was due to an outlying participant with an extremely long color recognition latency (3222 ms; the mean of the remaining participants was 331 ms). Exclusion of this participant resulted in a significant and large effect, $t(16) = 2.93$, $p = .010$, $\eta_g^2 = .20$. In agreement with H9, color was recognized significantly earlier than the objects shown in the affective pictures. A supplementary t-test showed that, in addition, color was detected significantly earlier than affect was experienced, $t(16) = 6.26$, $p < .001$, $\eta_g^2 = .55$. This finding remained robust even if the outlier was included, $t(17) = 2.68$, $p = .016$, $\eta_g^2 = .17$.

Analysis for Pictures as a Random Factor

By using the medians of the object recognition, valence detection, and affect latencies aggregated across different pictures as the dependent variable, we implicitly treated items (picture identity) as a fixed factor in the ANOVAs (Clark, 1973). It could be argued that pictures should be treated as a random factor to be able to generalize the findings to other pictures similar to those included in the experiment. To address this concern, we also computed *minF'* tests, which approximate the F tests of an ANOVA with both participants and items as random effects. *minF'* can be easily computed from the results of a by-subjects and a by-item ANOVA (F1 and F2 analyses; see Clark, 1973; Raaijmakers et al., 1999; Raaijmakers, 2003). We would like to point out, however, that several statistical experts (e.g., Edgington & Onghena, 2007; Siemer & Joormann, 2003) have argued that treating items as a random factor is questionable if, as in the present case, the items are not randomly sampled from a defined population.

The *minF'* analyses confirmed the significant difference between object recognition and affect (H1), *minF'*(1.00, 21.63) = 25.52, $p < .001$; between object recognition and valence detection (H5), *minF'*(1.00, 18.16) = 14.93, $p = .001$; and between valence detection and affect (H6), *minF'*(1.00, 23.11) = 14.80, $p = .001$.

Analyses for Individual Pictures

To determine whether the overall differences between the latencies of object recognition, valence detection, and affect are also found for individual pictures, we averaged the two reaction times per picture of each participant and then computed the 20% trimmed means of the latency differences for each picture across participants.

Descriptively, the predicted pattern of latency differences (object recognition $\prec$ valence $\prec$ affect) was found for all 24 pictures. Dependent-sample Yuen-t-tests for trimmed means (used to account for outliers; Wilcox, 2016) revealed significant ($p <$ .05, two-tailed) latency differences between object recognition and affect onset in 19 of the 24 cases, between object recognition and valence detection in 20 cases, and between valence detection and affect in 12 cases. The exception for the affect-object recognition difference were five unpleasant pictures ($M = 486$ ms delay, $p \geq .051$); the exceptions for the valence detection-object recognition difference were one pleasant and three unpleasant pictures ($M = 101$ ms delay, $p \geq .071$), and the exceptions for the affect-valence detection delay were one pleasant and eleven unpleasant pictures ($M = 446$ ms delay, $p \geq .056$). However, given that the nonsignificant latency differences were still relatively large in absolute terms and the Yuen-t-tests were based on a *df* of only 11, the nonsignificant results can most likely be attributed to an insufficient power of these tests to detect smaller latency differences.

We also tested the correlational hypotheses H2 and H7 on the level of the 24 individual pictures, by computing the correlation between the latencies of object recognition and affect latencies for each picture. To protect against outliers, the Spearman rank correlation (ρ) was used. The median correlation was ρ = .25 (range = -.01–.60; 23 of 24 positive) for object recognition and affect (H2); ρ = .31 (range = -.17–.73; 22 positive) for object recognition and valence detection (H7a); and ρ = .36 (range = -.03–.62; 22 positive) for valence detection and affect (H7b). In each case, a quarter (6–7) of the correlations were > .469 and thus passed the .05 significance level (two-tailed) for the given sample size.

Experiment 2

Preliminary ANOVA

The first preliminary ANOVA included the factor task (with all five sRT tasks included) and the control variables task order (object recognition or affect first), valence order (pleasant or unpleasant pictures judged first) and blur group (picture set A or B blurred). In the second preliminary ANOVA, the task factor was restricted to object recognition, valence detection, and affect; and blurring (blurred or sharp pictures), coloring (normal-colored or false-colored pictures) and picture valence (pleasant vs. unpleasant pictures) were added to the list of factors. Task type, valence, blurring and coloring were within-subject factors, whereas valence order, task order and blur group were between subject-factors. Huynh-Feldt corrections of the degrees of freedoms were applied when appropriate; the *p*-values for the control factors were Holm-corrected. These ANOVAs revealed no significant main or interaction effects for the control variables valence order (first ANOVA: $F \leq 1.73$, $p \geq .993$; second ANOVA: $F \leq 3.18$, $p > .999$), task order (first ANOVA: $F \leq 4.19$, $p \geq .689$; second ANOVA: $F \leq 6.86$, $p \geq .968$) or blur group (first ANOVA: $F \leq 4.19$, $p \geq .689$; second ANOVA: $F \leq 3.73$, $p \geq .968$). However, there was a significant and large main effect of task in both ANOVAs (first ANOVA, $F(1.33, 41.38) = 72.57$, $p < .001$, $\eta^2_g = .61$; second ANOVA: $F(1.12, 34.82) = 28.17$, $p < .001$, $\eta^2_g = .20$).

In the second ANOVA, the experimental manipulations also yielded significant effects: blurring $F(1, 31) = 16.00$, $p = .045$, $\eta^2_g = .01$; and coloring $F(1, 31) = 76.53$, $p < .001$, $\eta^2_g = .03$. In addition, coloring was involved in a two-way-interaction with task, $F(1.27, 39.28) = 36.14$, $p < .001$, $\eta^2_g = .03$, as well as in a three-way-interaction with task and valence, $F(1.08, 33.58) = 15.42$, $p = .037$, $\eta^2_g = .02$.

Latencies of Color Recognition and Appearance Detection

Hypothesis H9 states that the latency of color recognition is shorter than or equal to the latency of object recognition. Supporting this hypothesis, a one-factorial ANOVA revealed a significant effect of task, $F(1, 38) = 48.60$, $p < .001$, $\eta^2_g = .38$. A supplementary ANOVA showed that color was also detected earlier than affect was experienced, $F(1, 38) = 60.23$, $p < .001$, $\eta^2_g = .44$. Supporting H10, stimulus appearance detection was faster than color recognition, $F(1, 38) = 8.77$, $p = .005$, $\eta^2_g = .05$.

Analyses for Pictures as a Random Factor

To test to which degree the findings hold up if picture identity is treated as a random factor, we conducted three 2 (task) × 2 (blurring) × 2 (coloring) × 2 (picture valence) *minF*' analyses (Clark, 1973) on the latency medians, one each for H1 (object recognition vs. affect), H5 (object recognition vs. valence detection), and H6 (valence detection vs. affect).

Replicating corresponding effects of the by-subject analysis for H1, there were main effects of task, *minF*'(1.00, 45.34) = 31.22, $p < .001$; blurring, *minF*'(1.00, 40.16) = 5.02, $p = .031$; and coloring, *minF*'(1.00, 57.12) = 37.40, $p < .001$, as well as interactions between task and coloring, *minF*'(1.00, 58.65) = 26.91, $p < .001$, valence and coloring, *minF*'(1.00, 52.34) = 11.60, $p = .001$, and a significant three-way interaction comprising all three factors, *minF*'(1.00, 57.70) = 13.50, $p <$ .001. The main effect of task supports H1 for stimuli considered as a random factor. Furthermore, simple effects analyses confirmed that affect was later than object recognition at each of the four picture valence × coloring combinations, *minF*'s ≥ 6.25, *p*s ≤ .016 (H1).

The findings concerning Hypothesis H3a (blurring the pictures delays both object recognition and affect) were however only partially replicated by the simple effects *minF*'-tests; blurring significantly delayed affect, *minF*' (1.00, 39.08) = 5.23, $p =$ .027, but not object recognition *minF*'(1.00, 39.08) = 1.28, $p = .265$.

In agreement with H4a, simple effects ANOVAs conducted for the four task × picture valence combinations revealed that coloring significantly delayed the onset of affect, *minF*'(1.00, 49.00) = 25.18, $p < .001$ for the pleasant pictures, whereas the effect for unpleasant pictures was now only marginally significantly, *minF*(1.00, 47.79) = 3.44, $p = .070$. Also replicating the by-subjects analysis, the *minF*' test for coloring was not significant for object recognition, neither for the pleasant, *minF*(1.00, 41.90) = 1.52, $p = .225$, nor for the unpleasant pictures, *minF*(1.00, 24.77) = 0.86, $p = .363$.

The findings of the by-subject analysis for H5 were also largely replicated in the *minF*' analysis. There were significant main effects of task, *minF*'(1.00, 39.88) = 22.03, $p < .001$, and coloring, *F*(1.00, 58.52) = 13.46, $p = .001$; as well as significant interactions between valence and coloring, *minF*'(1.00, 57.75) = 5.85, $p = .019$, and valence and task, *minF*'(1.00, 44.82) = 11.02, $p = .002$, plus a significant three-way interaction comprising all three factors, *minF*'(1.00, 49.07) = 10.84, $p = .002$. The interaction between task and coloring was only marginally significant in the *minF*' analysis, *minF*'(1.00, 59.60) = 3.32, $p = .074$. In contrast to the by-subjects analysis, the main effect of blurring did not reach significance, *minF*'(1.00, 31.96) = 1.11, p = .300. Simple effects *minF*'-test confirmed that valence detection occurred later than object recognition at every coloring × picture valence combination, *minF*'s ≥ 5.13, *p*s < .029, except for the unpleasant, false-colored pictures, *minF*' = 1.06, $p =$ .310. This pattern was identical to that of the by-subject analysis.

As to H6, we obtained significant main effects of task, *minF*'(1.00, 49.55) = 22.74, $p < .001$, coloring, *minF*'(1.00, 59.10) = 38.01, $p < .001$, whereas the main effect of blurring was only marginally significant, *minF*'(1.00, 42.42) = 4.01, $p = .052$. In addition, there were again significant interaction effects between task and coloring, *minF*'(1.00, 56.03) = 23.07, $p < .001$, valence and coloring, *minF*'(1.00, 55.92) = 15.53, $p < .001$, and a significant three-way interaction comprising all three factors, *minF*'(1.00, 54.85) = 9.21, $p = .004$. Furthermore, we found a significant interaction effect of task and blurring, *minF*'(1.00, 54.81) = 5.92, $p = .018$. Simple-effects *minF*'-tests confirmed that valence detection was followed by affect for all coloring × picture valence combinations, *minF*'s ≥ 4.60, *p*s ≤ .037, and for both sharp and blurred pictures, *minF*'s ≥ 56.12, *p*s ≤ .001.

Finally, regarding hypothesis H8 (blurring and false-coloring both delay the detection of valence), the 2 (blurring) × 2 (coloring) × 2 (picture valence) *minF*'-test on the valence judgments confirmed the effects obtained in the by-subjects analysis: main effects of coloring, *minF*'(1.00, 59.12) = 10.71, $p = .002$, picture valence, *minF*'(1.00, 52.66) = 10.77, $p = .002$, and an interaction between coloring and picture valence, *minF*'(1.00, 57.88) = 13.61, $p \leq .001$; no other effects reached significance. Follow-up *t*-tests showed that coloring prolonged the latency of valence detection for pleasant pictures, *minF*'(1.00, 46.29) = 13.93, p = .001, but not for unpleasant pictures, *minF*'(1.00, 48.29) = 0.05, p = .833. Paralleling the by-subject analysis, then, H8 received only partial support.

In sum, the *minF*' analyses revealed that the findings of the ANOVAs were largely replicated if not only participants but also pictures were treated as random factors.

Analyses for Individual Pictures

To check whether the findings of the analyses for individual pictures conducted in Experiment 1 can be replicated, we repeated these analyses in Experiment 2 for the sharp, normal-colored pictures.

Regarding first the latency differences for object recognition, affect and valence detection, we averaged the two reaction times per picture of each participant for sharp, normal-colored pictures and then computed the 20% trimmed means of the latency differences for each picture across participants. Descriptively, the predicted positive latency difference between object recognition and valence was found for all pictures (pleasant: M_{Diff} = 194, SD_{Diff} = 76, range$_{\text{Diff}}$ = 128–413 ms, unpleasant: M_{Diff} = 140, SD_{Diff} = 49, range$_{\text{Diff}}$ = 67–218 ms). Dependent-sample Yuen-t-tests for trimmed means (Wilcox, 2016; Yuen, 1974) revealed that the latency differences reached significance for 11 of the 12 pleasant ($p < .025$, two-tailed) and for 10 of the 12 unpleasant ($p < .045$, two-tailed) pictures.

The difference between the latencies of affect and object recognition was also positive for all pictures (pleasant: M_{Diff} = 378, SD_{Diff} = 158, range$_{\text{Diff}}$ = 234–836 ms, unpleasant: M_{Diff} = 174, SD_{Diff} = 114, range$_{\text{Diff}}$ = 19–364 ms). These differences reached significance for all of the pleasant ($p < .020$, two-tailed) and for 5 of the 12 unpleasant pictures ($p < .049$, two-tailed).

Finally, the difference between affect and valence detection was positive for all pleasant and for 8 of the 12 unpleasant pictures (pleasant: M_{Diff} = 197, SD_{Diff} = 120, range$_{\text{Diff}}$ = 45–443 ms, unpleasant: M_{Diff} = 67, SD_{Diff} = 103, range$_{\text{Diff}}$ = -56–189 ms). The differences reached significance for 6 of the 12 pleasant pictures ($p < .039$, two-tailed), but in no case for the unpleasant pictures.

The test of the correlational hypotheses H2 and H7 conducted in Experiment 1 for individual pictures was also repeated in Experiment 2, by computing the Spearman rank correlation (ρ) between the latencies of object recognition and affect latencies for each of the 24 normal-colored and sharp pictures. The median correlation between object recognition and affect (H2) was ρ = .35 (range = .11–.51) for the pleasant, and ρ = .39 (range = .14–.61) for the unpleasant pictures; between object recognition and valence detection (H7a), it was ρ = .42 (range = -.14–.72; 11 of 12 positive) for the pleasant and ρ = .25 (range = -.03–.61, 10 of 12 positive) for the unpleasant pictures; and for valence detection and affect (H7b), the correlations was ρ = .37 (range = -.01–.44; 11 of 12 positive) for the pleasant, and ρ = .43 (range = -.02–.74; 11 of 12 positive) for the unpleasant pictures.

In sum, the findings of Experiment 1 regarding the hypothesis tests on the level of individual picture were mostly replicated in Experiment 2.

Additional References

Edgington, E. S., & Onghena, P. (2007). *Randomization tests* (4th ed.). CRC press.

Raaijmakers, J. G. (2003). A further look at the "language-as-fixed-effect fallacy". *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 57*(3), 141. https://doi.org/10.1037/h0087421

Siemer, M., & Joormann, J. (2003). Assumptions and consequences of treating providers in therapy studies as fixed versus random effects: Reply to Crits-Christoph, Tu, and Gallop (2003) and Serlin, Wampold, and Levin (2003). *Psychological Methods, 8*(4), 535–544. https://doi.org/10.1037/1082-989x.8.4.535

Appendix D

Author Contributions

Publication A (Reisenzein & Franikowski, 2022)

RR designed Experiments A1a and A2, PF and RR designed Experiments A1b and A3. A group of student interns did laboratory assessment for Experiments A1a and A2 under supervision of RR, Annemarie Bierstedt (student research assistant) and PF did laboratory assessment for Experiment A1b, Geraldine Quint, Isabel Kuhlmann, and Magdalena Korda (student research assistants) did laboratory assessment for Experiment A3 under supervision of PF. PF preprocessed and analyzed the data under supervision of RR. Both authors contributed to the interpretation of the data for both experiments. RR and PF wrote the manuscript (first draft provided by RR).

Publication B (Franikowski et al., 2021)

LK, PF, and RR designed the first experiment, PF and RR designed the second experiment, LK did the laboratory assessment for both experiments under supervision of PF, PF preprocessed and analyzed the data under supervision of RR. All authors contributed to the interpretation of the data of the first experiment, PF and RR interpreted the data of the second experiment. PF and RR wrote the manuscript (first draft provided by PF).

Publication C (Franikowski & Reisenzein, 2022)

PF and RR designed both experiments, PF and Annemarie Bierstedt (student research assistant) did laboratory assessment for the first experiment, Isabel Kuhlmann (student research assistant) did laboratory assessment for the second experiment under supervision of PF. PF preprocessed and analyzed the data under supervision of RR. Both authors contributed to the interpretation of the data for both experiments. PF and RR wrote the manuscript (first draft provided by PF).

Prof. Dr. Rainer Reisenzein　　　　Philipp Franikowski

Appendix E

Eigenständigkeitserklärung

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

Philipp Franikowski

Appendix F

Curriculum Vitae

Dipl.-Psych. **Philipp Franikowski**
Date of birth: 13 September 1993 in Grevesmühlen, Germany

Employment

Since 04/2022	**Research Assistant** Department of Personality Psychology / Psychological Assessment (Vertr.-Prof. Dr. Nicolas Becker) University of Greifswald
11/2017–03/2022	**Research Assistant** Department of General Psychology II (Prof. Dr. Rainer Reisenzein) University of Greifswald
05/2020–11/2020	**Research Assistant** Department of Sustainability Science and Applied Geography (Prof. Dr. Susanne Stoll-Kleemann) University of Greifswald

Education

Since 07/2019	**PhD Student** University of Greifswald Thesis: *Semantic or Affective Primacy: Perceptual Latencies of Object Recognition and Affect Measured With Temporal Judgments and Speeded Reaction Times*
10/20212–10/2017	**Diploma in Psychology** University of Greifswald
2004–2012	**Abitur (A-Level Equivalent)** Grevesmühlen

Scientific Publications

Peer-Reviewed Journal Articles

Nicolai, S., Franikowski, P., & Stoll-Kleemann, S. (2022). Predicting pro-environmental intention and behavior based on justice sensitivity, moral disengagement, and moral emotions: Results of two quota-sampling surveys. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.914366

Stoll-Kleemann, S., Nicolai, S., & Franikowski, P. (2022). Exploring the moral challenges of confronting high-carbon-emitting behavior: The role of emotions and media coverage. *Sustainability, 14*(10), 5742. https://doi.org/10.3390/su14105742

Franikowski, P., & Reisenzein, R. (2022). On the Latency of object recognition and affect: Evidence from speeded reaction time tasks. *Journal of Experimental Psychology: General*. Advance online publication. https://doi.org/10.1037/xge0001244

Reisenzein, R., & Franikowski, P.* (2022). On the latency of object recognition and affect: Evidence from temporal order and simultaneity judgments. *Emotion*. Advance online publication. https://doi.org/10.1037/emo0000991 (*Shared first authorship)

Franikowski, P., Kriegeskorte, L.-S., & Reisenzein, R. (2021). Perceptual latencies of object recognition and affect measured with the rotating spot method: Chronometric evidence for semantic primacy. *Emotion, 21*(8), 1744–1759. https://doi.org/10.1037/emo0000991

Seidlein, A.-H., Bettin, H., Franikowski, P., & Salloch, S. (2020). Gamified E-learning in medical terminology: The TERMInator tool. *BMC Medical Education, 20*(1). https://doi.org/10.1186/s12909-020-02204-3

Conference Contributions

Franikowski, P., & Reisenzein, R. (2022). *Semantic or affective primacy? Perceptual latencies of stimulus recognition and affect.* Vortrag auf der Konferenz der International Society for Research on Emotion (ISRE) 2022. Los Angeles, CA, USA.

Franikowski, P., & Reisenzein, R. (2022). *Semantic or affective primacy? Perceptual latencies of stimulus recognition and affect for sounds, measured with the rotating spot method.* Vortrag auf der 64. Tagung experimentell arbeitender Psychologen (TeaP). Köln, Germany.

Reisenzein, R., & Franikowski, P. (2022). *Semantic or affective primacy? Latencies of object recognition and affect: A comparison of rotating spot and speeded reaction time measures.* Vortrag auf der 64. Tagung experimentell arbeitender Psychologen (TeaP). Köln, Germany.

Franikowski, P., & Stoll-Kleemann, S. (2021). *Predicting high-carbon behavior: Development and validation of a German scale to assess moral disengagement in high-carbon contexts.* Vortrag auf der International Conference on Environmental Psychology (ICEP) 2021. Syrakus, Italien.

Franikowski, P., & Reisenzein, R. (2021). *Perceptual latencies of object recognition and affect measured with the rotating spot method.* Vortrag auf der 63. Tagung

experimentell arbeitender Psychologen (TeaP). Ulm, Germany.
Reisenzein, R., & Franikowski, P. (2021). *The Subjective Timing of Object Recognition and Affect.* Vortrag auf der 63. Tagung experimentell arbeitender Psychologen (TeaP). Ulm, Germany.

Teaching Experience

WS 17/18–WS 21/22	Motivation und Lernen
SS 18–SS 21	Emotion
WS 18/19–WS 21/22	Empirisch-Methodisches Praktikum
SS 22	Testtheorie und Testkonstruktion
WS 22/23	Diagnostische Verfahren
WS 22/23	Emotionsforschung

Danksagung

Ich möchte an dieser Stelle die Gelegenheit nutzen, um allen Personen zu danken, die mich während der gesamten Promotionszeit unterstützt haben. Um alle namentlich zu erwähnen, die mich auf dem Weg begleitet haben, fehlt mir natürlich hier der Platz.

An oberster Stelle gebührt mein größter Dank selbstverständlich meinem Doktorvater, Professor Rainer Reisenzein, der mich bei diesem Projekt betreute und durchgehend mit größter Hingabe unterstützte. Ich danke Ihnen auch besonders für die Freiheiten, die Sie mir bei der Gestaltung meiner Aufgaben als wissenschaftlicher Mitarbeiter einräumten, sodass ich mich in Bezug auf meine Lehre und meine Qualifizierung frei entfalten konnte. Vielen Dank auch für die stets motivierenden Worte und die Geduld bei allen Schritten.

Viele haben vielleicht nur eine "gute Seele" für die Rückendeckung, aber ich hatte gleich drei: Danke an Antje Punkt, Evelyn Reichel und Sylvia Scholz für die immer großartige Unterstützung bei allen größeren und kleineren Verwaltungsakten und Organisationsbelangen, ohne die der wissenschaftliche Alltag vermutlich kaum zu bewältigen gewesen wäre.

Außerdem möchte ich mich bei den Hilfskräften bedanken, die mich bei der Erhebung unterstützt haben: Annemarie Bierstedt, Adele Cyris, Magdalena Korda, Isabel Kuhlmann und Geraldine Quint. Dabei ein ganz besonderer Dank an Lara-Sophie Kriegeskorte – dafür, dass Du meine erste und dabei so unglaublich unkomplizierte Bachelorstudentin warst und dabei noch so wichtige Beiträge für die zweite Publikation beigesteuert hast. Und natürlich auch bei den Versuchspersonen, die an den Experimenten teilgenommen haben und ohne die diese Arbeit gar nicht möglich gewesen wäre.

Darüber hinaus danke ich meinen Kolleg*innen – insbesondere Carmen Pütz-Bonnéhs für die reset-tenden Mittagspausen – und langjährigen Freund*innen – insbesondere Marie Bernier, Lou Frankenstein, Arezoo Mazloomian und Julia Stendel; danke für Eure stets offenen Ohren, Eure Freundschaft und überhaupt, dass es Euch in meinem Leben gibt. Insbesondere vielen Dank an Carmen und Lou für Eure formalen und stilistischen Anmerkungen zu der Dissertation.

Zuletzt möchte ich mich bei meinen Eltern bedanken – für all die Möglichkeiten, die Ihr mir mit Eurer bedingungslosen Unterstützung vom ersten Tag an eröffnet habt. Und zu guter Letzt natürlich bei Kilian Alexander Wietschel – für Deine unschätzbare Unterstützung in allen erdenklichen Lebenslagen und überhaupt die Tatsache, Dich in meinem Leben zu haben, gibt es gar nicht die Worte, um meinen Dank angemessen auszudrücken.