Title: Studying unconscious processing: towards a consensus on best practices

Authors: François Stockart[1]*, Maor Schreiber[2]*, Pietro Amerio[3], David Carmel[4], Axel Cleeremans[5,6], Leon Deouell[7], Zoltan Dienes[8], Patxi Elosegi[9], Surya Gayet[10], Alon Goldstein[11], Adelina-Mihaela Halchin[12], Guido Hesselmann[13], Ruth Kimchi[14], Dominique Lamy[2,15] Leyla Loued-Khenissi[16], Sascha Meyen[17], Nitzan Micher[2], Michael Pitts[18], Roy Salomon[19], Kristian Sandberg[20], Iris A. Schnepf[21], Aaron Schurger[22,23], David R. Shanks[24], David Soto[25,26], Amir Tal[2], Darinka Trübutschek[27], Miguel A. Vadillo[28], Simon van Gaal[29], Itay Yaron[14], Zefan Zheng[30], Nathan Faivre[1]*, Liad Mudrik[2,6,15]*

Affiliations:
1. Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France
2. School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel
3. Consciousness, Cognition & Computation Group, Center for Research in Cognition & Neurosciences, ULB Neuroscience Institute, Université libre de Bruxelles, Bruxelles, Belgium
4. School of Psychology, Victoria University of Wellington, New Zealand
5. Center for Research in Cognition and Neuroscience, ULB Institute of Neuro-science, Universite libre de Bruxelles, Brussels, Belgium
6. Canadian Institute for Advanced Research (CIFAR), Brain, Mind, and Consciousness Program, Toronto, ON, Canada
7. Edmond and Lily Safra Center for Brain Sciences and Department of Psychology, The Hebrew University of Jerusalem, Israel
8. School of Psychology, University of Sussex, UK
9. Basque Center on Cognition, Brain and Language, San Sebastian, Spain
10. Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, The Netherlands
11. Xoltar Inc, Connecticut, USA
12. Cardiff University School of Psychology, Cardiff, UK
13. Psychologische Hochschule Berlin (PHB), Berlin, Germany
14. Dept. of Psychology & Institute of Information Processing and Decision Making, University of Haifa
15. Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel
16. Département des Neurosciences Cliniques, Lausanne University Hospital, Lausanne, Switzerland
17. University of Tübingen, Tübingen, Germany
18. Department of Psychology, Reed College, Portland, OR, USA
19. Dept. of Cognitive Sciences, University of Haifa
20. Center of Functionally Integrative Neuroscience, Aarhus University, Aarhus, Denmark.
21. Experimental Cognitive Science Group, Department of Computer Science, University of Tübingen, Tübingen, Germany
22. Institute for Interdisciplinary Brain and Behavioral Sciences, Chapman University, Orange, CA, USA

23. Department of Psychology, Crean College of Health and Behavioral Sciences, Chapman University, Orange, CA, USA
24. Division of Psychology & Language Sciences, University College London, England
25. Ikerbasque, Basque Foundation for Science, Bilbao, Spain
26. Basque Center on Cognition, Brain and Language, San Sebastian, Spain
27. Research Group Neural Circuits, Consciousness, and Cognition, Max Planck Institute for Empirical Aesthetics, Frankfurt/Main, Germany
28. Departamento de Psicología Básica, Universidad Autónoma de Madrid, Spain
29. Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
30. Research Group Neural Circuits, Consciousness, and Cognition, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

* Shared authorship

Corresponding author: Maor Schreiber

Abstract

The scope of unconscious processing has long been, and still remains, a hotly debated issue. This is driven in part by the current diversity of methods to manipulate and measure perceptual consciousness. We report 10 recommendations from a group of researchers representing a range of theoretical backgrounds, on how to design, analyze, and report the results of studies on unconscious processing. We also discuss a set of outstanding issues on which more research is necessary. The recommendations are envisioned as a set of "best practice" principles that serve as a guide for studying unconscious processing. While they are likely to change following theoretical and methodological development, at this point they are important for helping researchers make more informed decisions when designing experiments that test for unconscious processing, and thereby lead to a more convergent corpus of knowledge about the extent – and limits – of unconscious processing.

Keywords: Unconscious processing, perceptual awareness, awareness measures, recommendations, best practices

Introduction

The scope of **unconscious processing** (i.e., the effect of stimuli that are not consciously perceived on perceptual and cognitive processes, or on neural activity; see Glossary and further discussion below) is one of the most hotly contested topics in cognitive science (Eriksen, 1956; Marcel, 1974; Peirce & Jastrow, 1884; Peters et al., 2017; Shanks, 2017), and, accordingly, has been extensively researched (for reviews, see Axelrod, Bar, & Rees, 2015; Hassin, 2013; Kouider & Dehaene, 2007; Mudrik & Deouell, 2022; Newell & Shanks, 2014; Sterzer et al., 2014; Tsikandilakis et al., 2019). This research has often yielded contradictory results, fueling many longstanding and ongoing disagreements (Francken et al., 2022). While some argue that unconscious processing permeates even high-level cognitive functions (e.g., Bargh, 2014; Dijksterhuis & Nordgren, 2006; Hassin, 2013), such as executive control (van Gaal et al., 2012), working memory (Soto, Mantla, & Silvanto, 2011) and information integration (for reviews, see Mudrik et al., 2014; Hirschhorn et al., 2021), others claim it is more limited in scope and only encompasses overlearned, automatic, reflex-like processes (e.g., Chater, 2018; Hesselmann & Moors, 2015; Holender, 1986; Newell & Shanks, 2023; Searle, 1998).

A possible explanation for this state of affairs might stem from a lack of agreement on the way consciousness should be manipulated and measured. Researchers have employed various experimental procedures to suppress stimuli from awareness[1] (e.g., to render them invisible, see Breitmeyer, 2015; Kim & Blake, 2005; or inaudible, see Dykstra, Cariani, & Gutschalk, 2017), various different measures to assess participants' awareness of the stimuli (i.e., determine if - and to what extent - they consciously perceived the stimulus; Sandberg et al., 2010), and a range of approaches to analyze and interpret the data (Dienes et al., 1995; Fleming, 2017; Kristensen, Sandberg, & Bibby, 2020; Maniscalco & Lau 2012; Rausch et al., 2015; Sandberg et al., 2011; Schmidt & Vorberg, 2006; Shanks, 2017). This diversity, on its own, is not problematic as diverse methods leading to similar conclusions provide higher convergent validity (Campbell & Fiske, 1959). Methodological weaknesses of some of these procedures, however, might muddy the waters in the quest towards accurately characterizing unconscious processing.

Indeed, some have suggested that the field suffers from a lack of guidelines for "best practices" (Rothkirch & Hesselmann, 2017) and a clear definition of what is meant by 'unconscious processing' (Francken et al., 2022). Providing standards for how to conduct, analyze, and report

[1] In this paper, we use 'awareness' and 'consciousness' interchangeably.

findings of unconscious processing will help us establish a more informed body of knowledge, whereby studies can be more easily compared and integrated into a cohesive account. Such an endeavor will hopefully also help researchers steer clear of methodological pitfalls and minimize the number of findings that do not replicate (e.g., Biderman & Mudrik, 2018; Moors et al., 2016; Moors & Hesselmann, 2018; Lanfranco, Rabagliati, & Carmel, 2023; Meyen et al., 2022; Stein, Utz, & Van Opstal, 2020; Rabagliati, Robertson, & Carmel, 2018; Skora et al., 2023). Recommendations on how to measure awareness (Overgaard & Sandberg, 2012; Rausch & Zehetleitner, 2016), how to conduct response priming experiments (Schmidt, Haberkamp, & Schmidt, 2011), and how to evaluate claims of high-level unconscious processing (Moors et al., 2019) have been put forward before. While relevant, these recommendations focused on specific points and reflected the personal opinions and conclusions of their authors. Moreover, recently raised issues still await discussion (e.g., Meyen et al., 2022; Shanks, 2017; see further discussion below). The current work aims to provide a broader framework, integrating several different points of view and addressing prominent issues. Moreover, while a plurality of theoretical views and experimental approaches is desirable, a quantification of the state of agreement is useful to identify best practices (Beatty & Moore, 2010).

Here, we follow other fields where best practices have been established, such as EEG (Pernet et al., 2020; Picton et al., 2000) and fMRI research (Poldrack et al., 2017), and report an initiative aimed at developing guidelines to study unconscious processing. The following set of practical recommendations was established by a group of 34 researchers from different laboratories around the world, representing a range of theoretical backgrounds and views on consciousness. The recommendations were discussed in a series of three meetings employing consensus methods (Waggoner, Carline, & Durning, 2016), where key methodological choices were presented, discussed, and voted upon (see *Supplementary Materials: The process by which the recommendations were established)*. The meetings were followed by extensive iterations between the participants in the writing process, until a reasonable agreement was obtained around key claims and arguments. The guidelines presented here focus on recommendations that gained a satisfactory level of agreement in the voting procedure (at least two thirds). Other, more contentious issues are introduced as outstanding questions, with a range of possible solutions.

Notably, these guidelines do not target the study of consciousness in general. Instead, they focus on the challenge of demonstrating unconscious processing. This challenge is twofold; on the one hand, researchers should make sure that participants are indeed unaware of the stimuli, thereby overcoming ***contamination by conscious processing*.** On the other hand, they should present the stimuli in a way that allows them to evoke *sufficiently strong signals***,** to yield measurable effects (e.g., on behavioral measures of perception and cognition). Though many paradigms are aimed at entirely suppressing stimuli from awareness (for reviews, see Breitmeyer, 2005; Kim & Blake, 2005), achieving both goals is far from trivial (Michel, 2023; Reingold & Merikle, 1988; Seth et al., 2008), as we discuss below. Here, all recommendations are discussed in the context of behavioral studies, although some of them may be generalized to neuroimaging and electrophysiological research.

Our intention is to highlight the different dilemmas that should be considered when designing, conducting, and reporting the results of studies addressing unconscious processing, and to suggest possible solutions. There is no perfect recipe, of course. Each decision made by researchers has its advantages and disadvantages, and prioritizing one aspect (e.g., adding more measures of awareness to improve sensitivity) will lead to a cost in another (e.g., diluting participants' resources due to task load). Thus, our recommendations here should not be regarded as binding rules. Nevertheless, we encourage researchers to explicitly justify methodological decisions, to allow readers to estimate how they might have affected the obtained results. Future developments in methodologies and approaches may eventually lead to modifications of these guidelines. Yet in the present moment, we argue that adopting these guidelines as a reference point will allow the field to strengthen the rigor of its research methods and will direct researchers to key considerations that should be taken into account when devising new experiments.

Below, we first offer a definition of some of the key concepts used in unconscious processing research. We then present 10 recommendations for best practices (see also Box 1 for more general recommendations for improving the quality of experimental methods that also apply, and are especially important, to studying unconscious processing). Finally, we discuss some outstanding issues and possible solutions.

Unconscious processing: definition and measures

The word 'unconscious' has been used in the context of a multitude of phenomena, from subliminal effects through implicit processing to social biases and stereotypes (e.g., Gawronski, Hofmann, & Wilbur, 2006; Graham & Lowery, 2004). Here, we focus on evidence for

unconscious processing, which we define as the effect of stimulus features that are not consciously perceived on perceptual and cognitive processes, or on neural activity. Thus, our definition does not include cases of ‘implicit’ processing, where the stimuli themselves are consciously perceived, but participants are not aware of their effect (for reviews, see Dai et al, 2023; Giménez-Fernández et al., 2023; Moran, Nudler, & Bar-Anan, 2023; Shanks & St. John, 1994; Williams, 2020). Nevertheless, several of the issues considered here, such as the relative advantages of subjective and objective awareness measures, can also be applied to implicit processing research.

For a finding to be considered evidence for unconscious processing, two criteria should be fulfilled. First, *the* ***critical stimulus*** *(or stimulus feature) of interest should not be consciously perceived*. In typical experiments on unconscious processing, this is achieved in one of two ways: critical stimuli are presented (1) in a ’threshold’ or ‘liminal’ condition, where they are sometimes experienced and sometimes not, or (2) in a ‘subliminal’ condition, where they are never experienced (but in practice are sometimes consciously experienced in a fraction of the trials). In the latter case researchers try to present the stimuli such that they are just under the perceptual threshold, to promote sufficient neural activation to affect behavior. Whether the critical stimulus was indeed not consciously perceived is typically demonstrated using one or several ***direct measures of awareness***, examining whether participants report having no perception of it (or some aspect of it), and/or are unable to perform some explicit judgment that would be straightforward if the stimulus were consciously perceived. Second, *there should be evidence that the critical stimulus was nevertheless processed (i.e., a* ***processing measure****).* Such evidence is usually obtained by ***indirect measures*** *of unconscious processing*, which show processing of the critical stimulus at the behavioral or neural level.[2] For example, in a typical masked priming experiment (see **Figure 1** for more details), the direct measure of awareness consists of either detecting the presence of the critical stimulus - the prime - or detecting/discriminating one of its features; and the indirect measure of unconscious processing involves providing a behavioral response to a consciously perceived target stimulus. This response could be influenced by the critical stimulus.

Measures of awareness can be objective or subjective. With **objective measures,** participants are typically asked to make some judgment on the critical stimulus (e.g., “Was the stimulus tilted

[2] Note that in some cases, direct measures can also be used to infer unconscious processing, as we explain below.

to the left or to the right", see Parkes et al., 2001; or "was the stimulus present or absent"; Naccache & Dehaene, 2001). Objective measures are typically taken using a forced-choice question so the accuracy of participants' answers can be assessed. The critical stimulus is deemed not consciously perceived if objective task performance is not better than chance-level performance across trials (but see again footnote 2), or by demonstrating **null sensitivity** in the task (for example, using Signal Detection Theory measures; Green & Swets, 1966).

With **subjective measures,** participants report some aspects of the conscious experience triggered by the critical stimulus (e.g., clarity, phenomenal magnitude, confidence). Therefore, these reports cannot be considered as correct or incorrect, as they reflect the content of participants' own perceptual experiences. Using subjective measures, the critical stimulus is considered not consciously perceived if the participant reports no experience of the stimulus (Dienes, 2007). Yet another innovative approach formulates the subjective awareness measure as a 2-Interval Forced Choice (2IFC) question, allowing researchers to also conduct a sensitivity analysis on subjective responses (Elosegi, Mei & Soto, 2024; Amerio et al., 2024; Peters & Lau, 2015; see also De Gardelle & Mamassian, 2014). In this approach, participants are asked to report in which of two intervals they experienced, or more strongly experienced, the critical stimulus or its relevant feature. That way, although participants still report their subjective experience, one could compare their report to the veridical presentation schedule.

Both subjective and objective awareness measures have been criticized (for further discussion, see Hesselmann, 2013; Overgaard, 2015; Seth et al., 2008; Timmermans & Cleeremans, 2015; and the sections below). Generally speaking, performance on tasks that are not based on a forced-choice decision is potentially endangered by the **criterion problem** (Eriksen, 1960; Lloyd et al., 2013; see also Francken et al., 2022, Figure 2)[3]. To explain the criterion problem, we focus on non-forced choice subjective measures. There, given the same borderline (threshold) perceptual experience, one participant may report perceiving the stimulus while a more conservative participant would report not perceiving it (Schmidt, 2015; for a recent study demonstrating how this could affect the results, see Fahrenfort et al. 2024). Subjective measures may accordingly overestimate or underestimate unconscious processing, depending on how conservative or liberal, respectively, the participant's criterion is. Both conservative and liberal criteria for answers on subjective awareness measures are concerning; on the one hand,

[3] This is true for most subjective measures, and for objective detection tasks; Unless signal detection measures are used to measure sensitivity independently from criterion-placement (bias).

conservative criteria could potentially lead to unwarranted conclusions regarding the scope of unconscious processing, in case participants say they did not perceive the critical stimulus, when in fact they did have some faint perception of it. Conversely, liberal criteria could potentially lead to a reduced number of trials that are considered unconscious. Given the criterion problem, some researchers hold that the subjective awareness measure should be accompanied by chance-level performance in an objective measure (Andersen, Overgaard, & Tong, 2019; Holender, 1986; Hesselmann, Hebart, & Malach, 2011; Izatt et al., 2014; Merikle & Reingold, 1990).

Objective awareness measures suffer from different limitations. They may be too strict, in several ways: first, in order to ensure that objective performance is at chance level, researchers may choose to suppress the critical stimulus so strongly that it produces signals too weak to elicit measurable (unconscious) effects. As a result, unconscious processing may be underestimated (e.g., Michel, 2023; Reingold & Merikle, 1988). Second, above-chance discrimination performance in the objective awareness measure might indicate conscious processing, but can also be driven by unconscious processing (Merikle, 1984; Merikle, Smilek, & Eastwood, 2001; Michel, 2023; Schmidt & Vorberg, 2006). In fact, some researchers consider above-chance objective performance as a marker of *unconscious* processing, as long as participants subjectively deny consciously perceiving the stimulus (e.g., Lau & Passingham, 2006; Soto, Mäntylä, & Silvanto, 2011; Trübutschek et al., 2017; 2019a; 2019b). If this is indeed the case, using above-chance level objective performance as a marker of conscious processing may lead us to falsely conclude that participants were aware of the critical stimulus, despite them stating otherwise.

On the other hand, objective awareness measures might also lead to an *over*estimation of unconscious processing, depending on when they are taken, and on the nature of the task. If objective awareness measures are taken at the end of the experiment, participants might disengage due to a lack of motivation or due to fatigue, underestimating their level of awareness earlier in the experiment (Pratte & Rouder, 2009). If objective awareness measures are interleaved during the experiment, carryover effects may occur from trials in which the stimulus was consciously perceived (e.g., Lin & Murray, 2014). Another potential issue might occur if the feature of interest is very complex. In that case, having the objective awareness measure focus on that feature could again overestimate unconscious processing, as the objective task might be too difficult (e.g., determine if a scene includes an incongruent object, or whether a face is male or female). In that case, performance might be at chance even when participants consciously experienced some aspects of the stimulus (e.g., saw some parts of the scene, or the eyes of the face). Then, the stimuli will be classified as “not consciously perceived” although they were, to some extent. This partial conscious experience might explain an effect of interest obtained with the processing measure (e.g., Gelbard-Sagiv et al., 2016). Thus, objective awareness measures come with their own challenges, but provide complementary information to subjective measures.

Though the above definitions and approaches are still being discussed and shaped in the field (Michel, 2023; Rothkirch & Hesselmann, 2017), we adopt them here as a basis for the suggested recommendations, which we present next. These recommendations emerged as consensus items, based on the procedure described in more detail in the *Supplementary Material: The process by which the recommendations were established*.

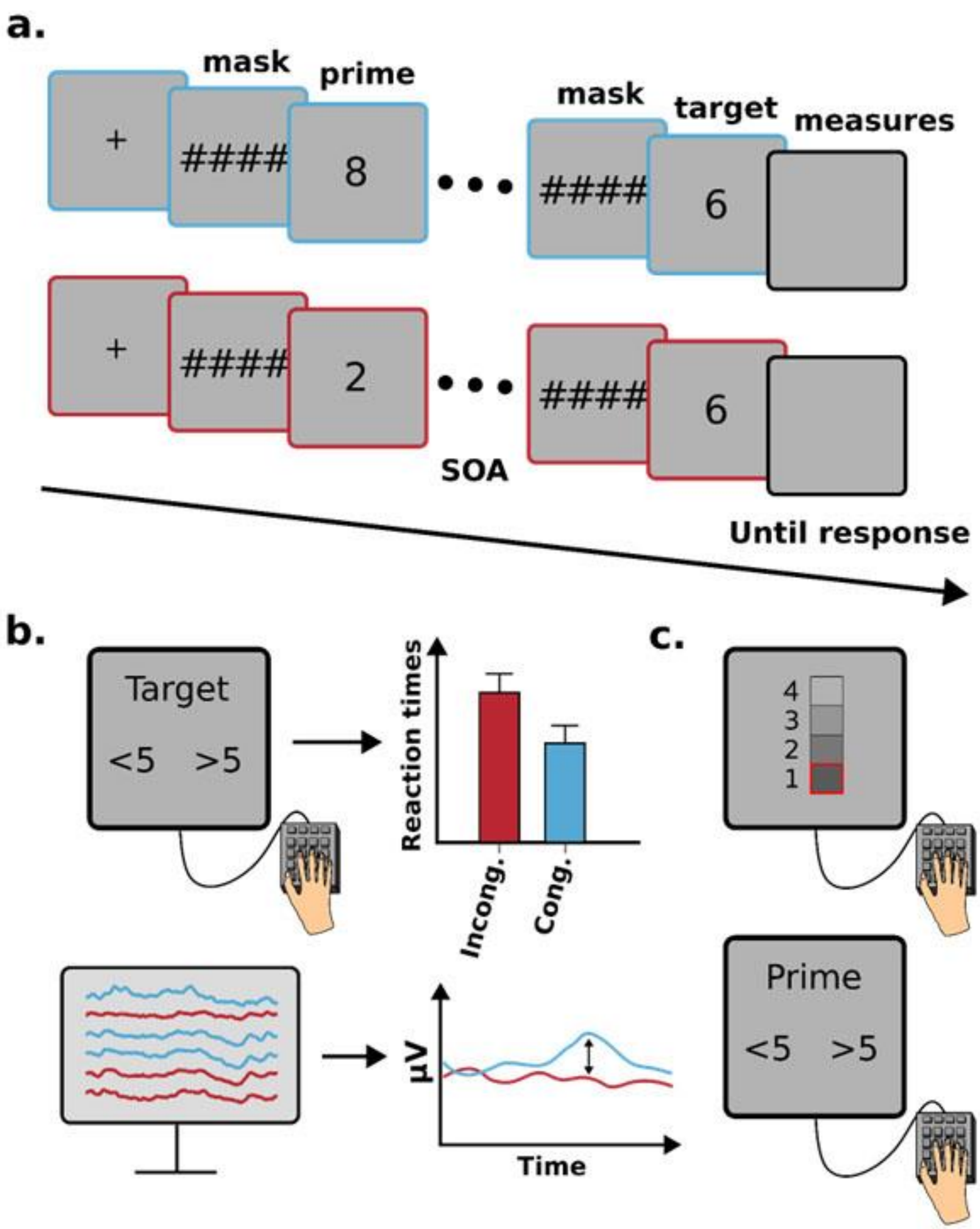

Figure 1. *A.* Example of a visual response priming task (Naccache & Dehaene, 2001). The task is to press one button if the target is larger than 5 and another if it is smaller than 5. The prime frame is preceded and followed by mask frames, and the stimulus onset asynchrony (SOA) determines the extent to which the prime is perceived. In congruent trials (blue), the prime and target both correspond to the same response

as both are larger than 5, while they correspond to opposite responses in incongruent trials (red). *B.* Processing measures can be either behavioral (top panels, showing slower reaction time on incongruent trials) or neural (bottom panels, schematically showing differential responses based on congruency), and consist of averaging responses or activity across trials in the congruent and incongruent conditions (right panels). *C.* Awareness measures can be subjective (for example, the Perceptual Awareness Scale in the top panel; see Recommendation 4 below) or objective (alternative forced choice on the prime - or the critical stimulus - in the bottom panel; here, determining whether the prime is smaller or larger than 5).

Recommendations

A: DESIGNING THE EXPERIMENT

1. Ensure adequate precision of the processing and awareness measures

To demonstrate unconscious processing, the sensitivity of both the measure of processing and the measure(s) of awareness must be sufficiently high to support robust inferences. Accordingly, researchers should allocate sufficient trials to both measures (Bates, Dufek, & Davis, 1992; Dufek, Bates, & Davis, 1995). This is especially important for the objective awareness measure, where low sensitivity might lead to a false negative conclusion, namely failing to find evidence for above-chance performance despite residual conscious processing. This could lead to an erroneous conclusion that the effect found on the processing measure can indeed be ascribed to unconscious processing (Leganes-Fonteneau et al., 2021; Meyen et al., 2022; Shanks, 2017; see further discussion in Recommendation 8). Yet, many studies in the field assign more trials to the processing measure than to the objective awareness measure, especially if the latter is administered in a later session (i.e., ‘a visibility post-test’; Vadillo et al., 2023). For example, in a recent reanalysis of data from 12 studies that included an objective awareness measure, the average number of trials in this measure was 82.09 (*SD* = 30.28, median = 76.91, IQR = 16.97), while the average number of trials in the processing measure was 334.79 (*SD* = 284.27, median = 212.77, IQR = 552.84) (Yaron et al, 2023).

How many trials are “enough”? This will inevitably depend on the specific combination of stimulation and task parameters, but we recommend that researchers provide a clear rationale for their choice in the methods section. The rationale can rely on simulations that take into account the estimated level of measurement noise, which differs across experimental situations (e.g., Yaron et al., 2023) and across objective and subjective awareness measures. Notably, in some experimental contexts, there might be an upper bound on the number of trials that can be presented, either because of the nature of the effect under investigation (e.g., inattentional blindness; Mack & Rock, 1998), or because of technical or experimental constraints (e.g., fatigue, training effects, time constraints in a neuroimaging facility or elsewhere). In these cases, researchers may consider running complementary control tests on the same participants (e.g., by asking participants in a neuroimaging experiment to perform a behavioral test in a mock scanner). When such solutions are not possible, we encourage researchers to acknowledge the limited sensitivity or precision of the awareness measure in the publication and be explicit about the level of evidence that they can provide for lack of conscious processing.

Importantly, some researchers use subjective measures to exclude trials where participants reported perceiving the stimuli. Such exclusion inevitably affects the number of trials in the processing measure and in the objective awareness measures. This again raises the issue of sensitivity, as well as potentially leading to a further complication, as some participants might have too few trials in one or several analyses (see Outstanding issue 6). Alongside these issues, post-hoc selection of trials evokes additional problems, which we discuss below (see Recommendation 8).

2. Use both subjective and objective awareness measures

A majority of participating researchers agreed that both subjective and objective awareness measures should be collected to assess awareness. Currently, various approaches are used in the field, with some studies using only a subjective awareness measure (e.g., Avneon & Lamy, 2019; Bahrami et al., 2010; Sperandio, Bond & Binda, 2018; Scott et al., 2018; Yang et al., 2017), some using only an objective one (e.g., Bruno et al., 2020; Filippova & Kostina, 2020; Haase & Fisk, 2020; Pesciarelli, Scorolli, & Cacciari, 2019; Stein, Utz, & Van Opstal, 2020; Tu et al., 2019), some using both (e.g., Biderman & Mudrik, 2018; Biderman et al., 2020; Faivre et al. 2016; Gayet et al., 2020; Hesselman et al., 2011; Izatt et al., 2014; Lamy et al., 2009; Sand & Nilsson, 2017), and others including only post-experiment debriefing (e.g., Chng, Yap, & Goh, 2019; Fadardi & Bazzaz, 2011; Park & Donaldson, 2016; Wang & Forster, 2015). Such heterogeneity of awareness measures across studies makes it challenging to compare and interpret results (Rothkirch & Hesselmann, 2017). Accordingly, we recommend collecting both types of awareness measures wherever feasible. This will make it possible to better understand how the awareness measures relate to each other (e.g., Stein et al., 2021). Importantly, as we acknowledge above, the two awareness measures might disagree; for example, one might find above-chance performance in the objective awareness measure, although participants subjectively report not perceiving the stimulus. We do not think that this situation necessarily implies that participants were, in fact, aware of the critical stimulus (and simply used a too conservative criterion for the subjective measure). This is only one out of two possible interpretations. The other is that performance in the objective awareness measure was driven by unconscious processing, and currently there is no definitive way to determine which of the two is correct. In this case, we recommend that researchers are clear about the theoretical basis of their interpretations.

When experimental tasks do not permit the collection of both types of awareness measures (for example, in inattentional blindness paradigms; Pitts et al., 2012), it is advisable to label the processing as “subjectively invisible" (this could also be applied in the case above, where objective performance is above chance level) or “objectively indiscriminable” (see Stein et al., 2021, for such an approach), to clearly specify on what type of measure the claim is substantiated.

3. Subjective awareness measure: a preference towards the Perceptual Awareness Scale (PAS)

Over the years, different subjective measures have been developed, ranging from dichotomous aware/unaware judgments (e.g., Ro et al., 2009), through numerical scales (either visibility scales, e.g., Dienes et al., 1995; Kuhn & Dienes, 2006, or confidence scales, e.g., Dienes et al., 1995; Kuhn & Dienes, 2006; Rothkirch et al., 2012) to wagering tasks (Persaud, McLeod, & Cowey, 2007). As described above, some even present the subjective measures as a forced choice task, by asking participants to indicate, for example, in which of two intervals they felt more confident that a stimulus appeared (Peters & Lau, 2015). Though all these measures can be used to gauge conscious perception, a majority of researchers in this project (79%, see Figure 2) preferred the PAS (Ramsøy & Overgaard, 2004), echoing preferences in the field at large (Francken et al., 2022). Note though, that some researchers felt quite strongly that this should not be the measure of choice (e.g., Amerio et al., 2024; Peters & Lau, 2015; Rajananda et al., 2020), because it does not completely eliminate the criterion problem (Fahrenfort et al., 2024) and because of its ambiguity concerning the characterization of “brief glimpses” (Dienes & Seth, 2010; Michel, 2023 and Recommendation 7).

The PAS is an ordinal scale that is held to capture situations of null awareness, by providing nuanced intermediate levels that can be used whenever participants think they might have perceived something despite having far from a clear conscious percept (Sandberg et al., 2010; Wierzchoń et al., 2014; 2019; though see Tunney & Shanks, 2003, for work implying that binary ratings might be more sensitive). Another advantage of this scale is its qualitative descriptions, which are based on people’s responses when asked to put their experience into words (Ramsøy & Overgaard, 2004). Arguments have also been made that it constitutes a more appropriate awareness measure than subjective measures that target other aspects of experience (Sandberg et al., 2010), such as confidence (though see Dienes & Seth, 2010; Michel, 2023; Szczepanowski et al., 2013). The PAS could be applied either on the entire stimulus or on the feature of interest, as we discuss below (Outstanding issue 2).

4. Objective awareness measure: Use a forced choice discrimination task on the feature of interest

Different studies have used various types of objective awareness measures, including identification of the critical stimulus (i.e., naming the stimulus; e.g., Bengson & Hutchison, 2007; Brunellière & Frauenfelder, 2014), and performing a discrimination task that pertains to one of its features (e.g., larger or smaller than five as in Figure 1; Filippova & Kostina, 2020; Kouider & Dupoux, 2001; Lourenco, Ayzenberg, & Lyu, 2016; Naccache & Dehaene, 2001; Ortells et al., 2016). We recommend having participants perform a forced-choice discrimination task. Discrimination tasks rely on simple instructions, thereby minimizing confusion for the participants who often perform dual, or even triple tasks on each trial (i.e., the processing measure, the subjective and the objective awareness measures).

Importantly, the objective task can be focused on several aspects of the stimulus display, including whether the critical stimulus is present or absent, and features of the critical stimulus (e.g., is a grating tilted left or right). We recommend focusing the objective awareness measure on the feature of interest, i.e., the feature that is taken to influence the processing measure (Dulany, 1963; Michel, 2023; Schmidt & Biafora, 2024). This approach is advantageous because it ensures a high level of task similarity between the two measures, which may increase unconscious effects (Reingold & Merikle, 1988; Schmidt, Haberkamp, & Schmidt, 2011). It is also preferential over using a detection task on the entire stimulus, because the latter might underestimate unconscious processing by posing a too-stringent test (Michel, 2023). As an example, consider a priming task probing number processing (represented in Figure 1), where digits are presented as primes and targets, and participants are asked to determine whether the target digit is larger or smaller than 5. There, the unconsciously processed feature of interest is whether the prime digit is larger or smaller than 5, and so the objective awareness measure should pose that exact question (rather than, for example, reporting the specific digit). As another example, if one looks for differential neural activity between invisible emotional and non-emotional faces, the objective awareness measure should involve a discrimination task between an emotional and a neutral face[4].

A downside of probing the feature of interest that applies only to priming studies is that the response about the target is likely to overshadow the response about the critical stimulus (here, the prime; Peremen & Lamy, 2014). In the above example, if the prime was the digit 2 but the target was the digit 6, participants might have had some conscious experience of the prime, but confused it with the target, and erroneously reported that the prime was larger than 5. Such contamination by processing of the clearly visible target renders the objective awareness measure less sensitive and more prone to errors. This issue can be mitigated, to some extent, in several ways. First, by counterbalancing the order of the objective awareness measure and the processing measure (in experiments that use priming to probe unconscious processing; Stein & Peelen, 2021). Second, by adding trials within the main experiment where the target is not presented (Faivre & Kouider, 2011). However, if the target contributes to the invisibility of the critical stimulus, this might not reflect participants' processing of this stimulus in target-present trials. This too can be overcome by presenting a neutral target, rather than no target (e.g., in the example above, presenting a symbol instead of a digit), so that the target could still contribute to masking the critical stimulus, yet without evoking a competing representation. This too, unfortunately, evokes complications, as participants are required to task-switch between trials, such that on some trials they perform a task about the target and on other trials, they perform a task about the critical stimulus. This task-switching might require additional resources that would tamper with the unconscious process of interest. Thus, this requires some piloting, or a practice session, to make sure participants can indeed perform this task.

[4] Note that it is sometime more complicated. Consider, for example, a priming task probing global versus local amodal completion, where an occluded figure is presented as a prime and its amodal completion (global or local) as targets. In this case, the objective awareness measure cannot involve the critical stimulus (in this case the occluded prime), but rather, a discrimination between global and local completions of the prime (Kimchi et al., 2023).

5. Collect all measures on a trial-by-trial basis (when possible)

Objective and subjective awareness measures can either be administered on a trial-by-trial basis, that is, on the same trials on which the processing measure is obtained (e.g., Biderman & Mudrik, 2018; Sand & Nilsson, 2017), or separately, in an additional session (e.g., Bruno et al., 2020; Stein, Utz, & Van Opstal, 2020; Tu et al., 2019). We recommend the former approach, when possible, for two main reasons. First, measuring consciousness in a separate session might yield an inaccurate estimate due to differences between the main experimental session and the separate awareness measurement session. These differences might stem from order effects, like learning on the one hand, and fatigue on the other hand, though such differences can possibly be mitigated by using an ABAB design (see Outstanding issue 3). Perhaps more critically, the main experimental session and the separate awareness session also commonly involve differences between the tasks (Lin & Murray, 2014). These differences might affect findings in different ways. Participants might be less motivated and less engaged in a separate session, where the main task focuses on awareness of the critical stimulus, which is typically difficult to consciously perceive. Unmotivated participants may underperform on this task, leading to an underestimation of their conscious perception of the critical stimulus in the main session (Pratte & Rouder, 2009; Note that this motivational problem is not relevant for experiments using liminal stimulation, see Outstanding issue 3). Conversely, participants might direct more attention to the critical stimulus in the separate session, whereas in the main session, they had to provide a report on the target. This change in attention allocation might lead to an overestimation of the level of conscious processing in the main session.

The second reason why testing for awareness on a trial-by-trial basis is preferred is that the level of awareness most likely fluctuates from trial to trial (Mathewson et al., 2009), and this cannot be captured when testing awareness in a separate session. Thus, even if the separate awareness session indicates no awareness, it does not guarantee that there were no trials with conscious perception in the main session. Probing for awareness on a trial-by-trial basis allows the researcher to identify trials in which participants did not consciously perceive the stimulus (although such post-hoc selection poses the risk of regression to the mean; see Recommendation 8 below).

Although it is advantageous overall, the inclusion of trial-by-trial awareness measures also comes with disadvantages. The first was already discussed in Recommendation 4, with respect to priming studies, where the objective awareness measure might be contaminated by the visible target. A second disadvantage of including trial-by-trial awareness measures is that they might decrease the effects captured with processing measures due to shared resource allocation (i.e., dual task costs, see Kiefer, Harpaintner, Rohr, & Wentura, 2023; Fischer, Kiesel, Kunde, & Schubert, 2011; but see Biafora & Schmidt, 2022, where the trial-by-trial costs were mitigated by training). At the very least, participants will be presented with two tasks, or - in experiments where the processing measure involves an additional task (e.g., a priming experiment where a visible target is presented) - three tasks. For some research questions, adding many tasks after each stimulus may not be the optimal solution (see Jiang, Zhang, & van Gaal, 2015). This concern can be mitigated by a more thorough practice session at the beginning of the experiment.

Notably, in some cases it is impossible to administer the objective awareness measure during the task (e.g., in no-report paradigms, Frässle et al., 2014; Pitts et al., 2014). In these cases, we recommend administering an objective awareness measure in a separate session. The order in which the measures should be administered within a trial also remains controversial. These issues are discussed further in the Outstanding issues section.

6. Add catch trials in which no stimulus (or feature of interest) is presented

When using subjective measures, it is often helpful to add catch trials in which no stimulus is presented (e.g., Avneon & Lamy, 2019; Micher & Lamy, 2023) to better assess the validity of participants' subjective reports. Specifically, catch trials allow researchers to estimate the tendency to report awareness when no stimulus was presented. In case the subjective measure addresses awareness of a feature, rather than the presence of the critical stimulus (see Outstanding issue 2), catch trials should lack this feature (when possible). This could entail, for example, presenting a two-headed arrow or a horizontal line as a catch trial in a study that examines orientation processing of right/left arrows.

Catch trials may indicate that some participants report consciously perceiving stimuli even when none were presented, as is expected on any signal detection model. One option is to exclude these participants (keeping in mind, however, that reports of brief glimpses – as opposed to reports of clearly seeing the stimulus - are common even when no stimulus is presented; Sandberg et al., 2014). Because such an exclusion is driven by participants' general use of the scale and is orthogonal to the effect of interest, it should not evoke regression to the mean (see Recommendation 8). Another option is simply to report the number of such participants, and compare the results with and without them (taking into account that in some cases, participants might have false perceptions, but these are not directly relevant to the study of unconscious processing).

B. Analyzing the data

7. When using the PAS, trials rated as "brief glimpse" should not be classified as "unaware"

When using the PAS as a subjective awareness measure, a majority thought it best not to consider trials rated with the second lowest level of invisibility ("brief glimpse" trials) as trials lacking conscious perception. However, some of us thought otherwise, and the agreement scores were lowest amongst the recommendations we established (71%, see Figure 2). The rationale of the majority view was that if participants are instructed appropriately, this second lowest rating on the PAS should indicate that they consciously perceived some aspect of the stimulus to some extent (even if briefly). Under this view, it is advisable to only focus on trials where participants clearly stated not having perceived the stimulus ("no awareness" trials). Moreover, some of us noted that based on their experience, "brief glimpse" ratings are sometimes used by participants as a "default rating" when they do not pay attention to the stimulus, close their eyes, etc. (though this has not been systematically demonstrated), so that including these trials in the analysis might contaminate the effect of interest with noise and lead to underestimating unconscious processing (e.g., Avneon & Lamy, 2019; Kimchi et al., 2018; Ophir et al., 2020; Micher & Lamy, 2023). In contrast, the rationale of the minority group for considering such trials as "unaware" is that when saying that they only got a brief glimpse of the stimulus, participants report not consciously experiencing the feature of interest. Thus, not including these trials might lead to an underestimation of unconscious processing (Dienes & Seth, 2010; Michel, 2023). When deciding which PAS level is relevant to claims of subliminality, this decision should be justified according to what content is relevant to the task. In cases where there are sufficient trials, a possible approach is to provide results of analyses for each PAS level (see Outstanding issue 7, and also Ophir et al., 2020, for an analysis that better characterizes "brief glimpse" trials).

8. Examine the possible effect of misclassification due to measurement error

A common question that arises in studies on unconscious processing is whether to perform any type of post-hoc selection on the data. Two such selections are commonly used. First, excluding participants based on above-chance performance in the objective awareness measure (determined, for example, using a predefined threshold (e.g., >60%), a binomial test, or a Bayesian test; Rouder et al., 2007). Second, excluding trials based on the subjective awareness measure, keeping only trials where participants reported not consciously experiencing the critical stimulus.

The motivation for excluding participants who show above-chance performance, or trials where the stimulus was consciously perceived, is to ensure that the evidence for processing is not contaminated by such participants or trials, and genuinely reflects unconscious processing. However, *post-hoc* selection of data causes regression to the mean, which can in turn lead to false positive results (Schmidt, 2015; Shanks, 2017). To explain, measurement error in the awareness measure(s) (which is probably inevitable) means that some aware participants or trials are likely to be misclassified as unaware and vice versa. If awareness and processing performance are positively correlated, then for these misclassified participants or trials, their processing performance will regress towards the group mean and evidence of unconscious processing will wrongly be inferred. Indeed, evidence suggests that the reliability of many awareness measures is modest at best (Rothkirch, Shanks, & Hesselmann, 2022; Stein et al., 2024; Yaron et al., 2023), making such misclassification unavoidable. We deem it essential that studies of unconscious processing provide a convincing explanation of why their results are not contaminated by misclassification due to measurement error.

One might strive to avoid excluding participants/trials altogether, by designing the experiment using strong enough suppression methods or by applying individual calibration, so that each participant's objective performance is at chance-level and that all trials are rated as unaware. This might not be a realistic expectation, however: even if calibration succeeds, practice effects occurring over the course of the experiment might lower perceptual thresholds, and stimulus visibility might vary unpredictably on a trial-by-trial basis even given identical stimulation on each trial. In addition, using a suppression method that is too strong might excessively degrade the critical stimulus to the point that it cannot be processed at all. A more feasible solution is therefore to exclude participants or trials and then account for post-hoc data selection at the analysis stage. Several methods have been suggested for testing whether the obtained results might indeed be driven by regression to the mean (e.g., Dienes, 2024; Goldstein, Sklar, & Siegelman, 2022; Leganes-Fonteneau et al, 2018; Yaron et al., 2023; Rothkirch et al., 2022). These methods are still new and under development, so we refrain from recommending a specific one (see the *Supplementary Material: How to account for regression to the mean at the analysis stage?* for details). We suggest instead that researchers choose the method that seems most suitable to their experimental context.

In addition, to determine whether participant exclusion affected the overall results, we recommend that researchers also analyze the data of participants who perform above chance-level in the objective awareness measures (given that there are enough such participants), and report them (see further discussion in the reporting section below).

C. Reporting the results and writing the manuscript

9. Precisely define what is meant by 'unconscious processing'

The term 'unconscious processing' is often used with multiple interpretations, for instance in studies exploring the processing of suppressed stimuli (for review, see Mudrik & Deouell, 2022), experiments focusing on implicit processes affecting behavior (Greenwald & Banaji, 1995; Maresch et al., 2021), or studies of cognitive and social biases that affect people's decisions without them being aware of those biases (e.g., Gawronski, Hofmann, & Wilbur, 2006; Graham & Lowery, 2004). We accordingly recommend that researchers adopt a cautious approach and be precise about the meaning that they are ascribing to key terms (see Glossary for definitions of key terms).

More specifically, the following points should be made clear in any report of studies on unconscious processing: (a) what exactly participants were (presumed to be) unaware of (the entire stimulus, a specific feature of the stimulus, etc.); (b) how did the choice of the suppression method potentially affect the type of evoked unconscious processing (Breitmeyer, 2015; Dubois & Faivre, 2014; Peremen & Lamy, 2014); (c) how was awareness measured, i.e., via an objective awareness measure, a subjective awareness measure, or both; and (d) what were the instructions given to participants regarding the measures.

10. Justify the chosen measures and provide conditional interpretations

Given the great heterogeneity of methods in the field of consciousness studies, and the potential effect of methodological choices on the results (Yaron et al., 2022), we encourage researchers to explicitly justify their choices, and in particular the choice of awareness measures. This entails explaining what kind of results are held to demonstrate lack of awareness of the critical stimulus or stimulus feature (this seems especially important given the lack of consensus around definitions). It is also important to directly address criticisms of the chosen measures (discussed above). Hence, we further recommend that researchers explicitly state in the discussion section exactly how they expect their results to be affected by potential criticisms of their awareness measure(s). For example, this would involve stating how interpretations would change if there was indeed a bias in the subjective awareness measure or if the results were affected by poor reliability of these measures (e.g., misclassification of awareness due to measurement error). This approach ensures that claims are appropriately conditioned on the extent to which the critical stimuli were indeed successfully suppressed.

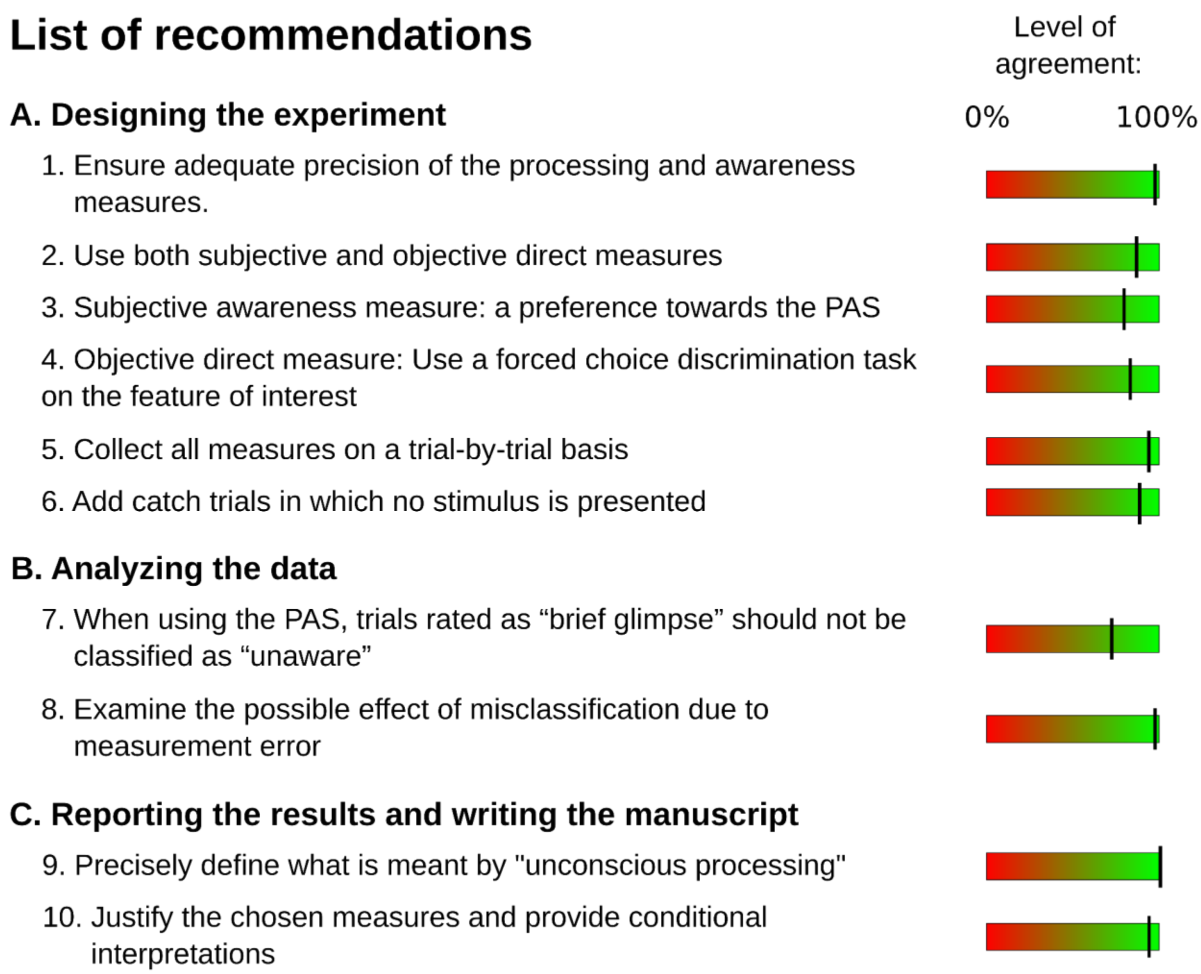

Figure 2. Ten recommendations for unconscious processing research. The level of agreement across the voters for each recommendation is shown on the right. The number of voters was 28[5] for all recommendations except for recommendations 6 and 9 where it was 24 and recommendation 10 where it was 20. For the original phrasing of the questions see *Supplementary Material: The original list of recommendations that were voted on during the first two meetings.*

Outstanding issues

Several points raised in our discussions did not yield a consensus, highlighting current controversies and ongoing challenges in unconscious processing research. Below, we elaborate on the main unresolved issues in the hope of accelerating attempts to resolve them. See a summary of these issues, possible solutions and their respective pros and cons, in Figures 3 (issues 1-4) and 4 (issues 5-8).

1. The order in which awareness measures are administered within a trial

In recommendations 2 and 5, we suggested administering three measures on each trial: the processing measure, and the subjective and objective awareness measures. When a separate task is needed for the processing measure, it is typically administered first, since many of these tasks require a speeded response. There is relatively little data on the effect of task ordering of the two awareness measures (e.g., Wierzchon et al., 2014). Thus, we chose not to recommend a specific order.

Possible options include having the objective awareness measure precede the subjective one, or vice versa, as well as alternating between the two (though this might evoke confusion). Another alternative is to probe for both awareness measures with a single question, thereby requiring a single response (e.g., Stein et al., 2021; the subjective-objective measure of awareness, Gelbard-Sagiv et al., 2016;) or the prime assessment grid, Sand & Nilsson, 2017). Besides shortening trial duration and reducing the chance of response priming between the processing and awareness measures (as a different motor response is required), one advantage of simultaneous probing is that it has been shown

[5] Although the process incorporated 34 researchers (two of whom decided to withdraw during the writing process), not all co-authors participated in all meetings where voting took place. Recommendation 6 was voted upon in a subsequent meeting and recommendations 9 and 10 were voted upon online.

to improve test-retest reliability in related research (Guggenmos, 2021; Vadillo et al., 2022). Also, it might reduce the response load, as participants only have two responses to provide on each trial, rather than three. On the other hand, it is also possible that participants find it more confusing to incorporate two responses within one task, and thus it may be more error-prone, in addition to being slower/less meaningful in the reaction time measure. Overall, more data should be acquired to better appreciate the advantages and disadvantages of each approach.

2. The feature to which subjective measures refer

Subjective measures can be used to probe the conscious perception of the presence of an entire stimulus, or of a specific feature of interest (e.g., its color, orientation, etc.). Although the former yields a stricter test, which might eliminate partial awareness of the critical stimulus and contamination of the reported effects by conscious processing (Kouider & Dupoux, 2004; Kouider et al., 2010; Gelbard-Sagiv et al., 2016), the latter tests what is relevant to the subliminal perception enabling task performance (Michel, 2023). Also, applying the subjective measures on the feature of interest is advantageous as it allows researchers to compare its results with that of the objective awareness measure, which, according to our recommendation, should refer to the feature of interest. For example, in a study investigating unconscious processing of facial expressions, participants might remain unaware of the expression even if they manage to get a glimpse of the hairline, or if they notice that something appeared without having any experience of its content, as typically occurs in visual crowding (Faivre & Kouider, 2011; Schmidt & Biafora, 2022). Under that approach, it seems advisable to use subjective measures to assess participants' awareness of the expression, rather than of the entire stimulus. Note that when using the PAS as a subjective measure, such usage might require modifying its verbal labels to the relevant context, which might require some verification process (cf. Sandberg & Overgaard, 2015). Importantly, the scale should be formulated so that the lowest level clearly means having no awareness whatsoever of either the feature of interest or the stimulus, depending on one's choice (Ramsøy & Overgaard, 2004; Peremen & Lamy, 2014).

3. Block ordering when the objective awareness measure is administered separately

Traditionally, separate awareness measurement sessions are administered at the beginning or the end of the experiment so as to minimize effects of training in the main session. However, this could evoke effects of task order, training and fatigue (see Recommendation 5). This can be mitigated by collecting these data both at the beginning and at the end of the experiment, or by using an ABAB design, where the main experiment (A) and the awareness test block (B) are administered interchangeably. Note that these approaches are not possible in studies wherein administering the objective awareness measure first could change the outcome of the processing measure (e.g., inattentional blindness). Also, this design does not address all the disadvantages of separating the different measures, which is why we generally recommend taking the objective awareness measure on a trial-by-trial basis during the main session (see Recommendation 5).

4. Including clearly-visible stimuli

Some studies include trials in which highly salient stimuli are intended to be clearly consciously perceived by participants, for example by presenting them at a higher contrast or not masking them. Arguably, these trials can be used to assess and improve the reliability of the objective awareness measure (Lin & Murray, 2014; Pratte & Rouder, 2009): when participants fail to perceive the critical stimulus on (almost) every trial, they may become demotivated and adopt a constant, meaningless response strategy. Including highly salient stimuli might alleviate this problem by ensuring that participants are more motivated, less bored and more alert (Pratte & Rouder, 2009). However, such trials can also lead to longer experiments, adding to the effects of fatigue and training that might influence the overall results. There is also evidence that including trials with clearly visible stimuli increases the visibility of stimuli in other trials, which would have not been consciously perceived otherwise (Lin & Murray, 2014), and that consciously exposing participants to the stimuli facilitates unconscious processing (to the point that it might not be present without such an exposure; Faivre et al., 2014; Gayet et al., 2014). Another potential concern is that including such trials will raise participants' criterion, making them report not seeing stimuli that have been faintly experienced.

By and large, these possible positive and negative effects of including trials with highly salient stimuli have not been tested empirically in a systematic manner. For example, the test-retest reliability of the objective awareness measure with or without the inclusion of these stimuli has not been compared. We accordingly encourage researchers to perform such tests and report their results. It is also important to investigate other methods for keeping motivation high while responding to the objective awareness measures. One avenue involves incentivizing participants to maximize their discrimination performance (e.g., via monetary payoffs). Participants may perform better when paid, compared to when they are not, with evidence that token payment is as effective as monetary payment (Johnson & Bickel, 2002). It will be relevant to test whether such effects can also be demonstrated in the context of objective awareness measures.

	1. The order in which awareness measures are administered within a trial	
	Sequentially	Simultaneously
Advantages	Reduced risk of errors and confusion	Shorter trial duration, potential reduction of cognitive load
Disdvantages	Longer time, order effects	Potentially confusing, response load

	2. The feature to which the subjective measures refers	
	The entire stimulus	The feature of interest
Advantages	Stricter test of awareness	Compatible with the other measures, reduced risk of underestimating unconscious processing
Disdvantages	Risk of underestimating unconscious processing, might create a mismatch with the indirect and objective direct measures, which typically refer to the feature of interest	Less strict, possible effects of partial awareness

	3. Block ordering when the objective awareness measure is administered separately		
	Post-test	Pre-test and post-test	Alternating
Advantages	Minimizes effects of training in the main task, so participants are kept unaware of the stimulus	Compared to a post-test only, minimizes effects of fatigue in the pre-test	Order effects can be measured
Disdvantages	Participants may be fatigued leading to underestimation of awareness	Lengthens experiments, training effects, participants might be more attentive to the stimulus in the main task	Awareness measures may yield inconsistent results between pre- and post-group

	4. Including clearly-visible stimuli
Advantages	Allows the assessment of the reliability of the direct measures, keeps participants motivated
Disdvantages	Lengthens experiments, might evoke a criterion shift or increase the visibility of other stimuli

Figure 3. A summary of outstanding issues 1-4.

5. Using the liminal or subliminal approach

Many studies in the field have taken the subliminal approach, aiming to present all critical stimuli under the detection threshold (e.g., Van den Bussche et al., 2009). To compare unconscious processing with conscious processing, some have compared the processing of such subliminal stimuli to that of supraliminal, and sometimes highly salient, ones (e.g., Kiefer & Spitzer, 2000). However, this approach introduces a confound whereby the conditions also differ in the physical properties of the stimuli. Accordingly, some researchers have opted for either using bistable stimuli, where perception fluctuates between two perceptual experiences despite the stimulus being constant (for reviews, see Rodríguez-Martínez & Castillo-Parra, 2018; Sterzer, Kleinschmidt, & Rees, 2009), or presenting stimuli around the threshold, so that the same stimuli evoke conscious or unconscious processing on different trials (Andersen et al., 2016; Avneon & Lamy, 2018; Kimchi et al., 2018; Micher & Lamy, 2023; Peremen & Lamy, 2014; Van den Bussche et al., 2013). These bistable perception and liminal-presentation approaches allow researchers to assess the extent to which different processes exclusively vary with conscious perception (Lamy, Ophir, & Avneon, 2019). Furthermore, they might also help avoid underestimating unconscious processing as the stimuli are less degraded and participants are more motivated (as also indicated in Outstanding issue 4). These approaches are based on post-hoc trial sorting, however, and therefore suffer from the shortcomings discussed above. First, since aware and unaware trials are classified based on the subjective awareness measure, the criterion problem should be addressed (Eriksen, 1960; Fahrenfort et al., 2024; Lloyd et al., 2013). A possible solution can be to also administer an objective awareness measure (see Recommendation 2), to validate participants' subjective ratings (Lamy et al., 2015). Second, trial selection might influence the results due to effects of misclassification induced by measurement error (Fahrenfort et al., 2024; Shanks, 2017; see Recommendation 8). Finally, as discussed above, conscious exposure to stimuli might affect the way they are processed without awareness (e.g., increasing the likelihood that they are unconsciously processed; Faivre et al., 2014). Future studies should focus on addressing the shortcomings associated with taking either the liminal or the subliminal approach.

6. Dealing with participants having too few trials due to trial exclusion

In many experiments targeting unconscious processing that use subjective measures, trials in which participants subjectively reported perceiving the stimuli are removed from analysis, which leads to some participants having too few trials in one or several analyses. There are at least two options to deal with this issue, and a combination of them may be used, depending on the experimental situation. The first is to either exclude participants with a low number of trials in at least one of the conditions of interest (this number should be decided on in advance). Because such an exclusion is driven by participants' general use of the scale and is orthogonal to the effect of interest, it should not evoke regression to the mean (see Recommendation 8). Another option is to report the number of such participants, and compare the results with and without them. Alternatively, one can compensate for the low number of trials using appropriate statistical methods (e.g., Generalized Linear Mixed Models; see Brysbaert & Stevens, 2018, for further discussion), or to minimize trial exclusions by designing an experiment in which stimulus intensity is calibrated for each participant at the beginning or during the experiment to be below the detection threshold. Notably, though, such a design is hard to achieve and might lead to excessive suppression of the stimuli, potentially underestimating unconscious processing. Also, it may not apply to all experimental contexts (e.g., the liminal paradigm; Lamy et al., 2019; or inattentional blindness Mack & Rock, 1998).

7. Performing multiverse analyses

In multiverse analyses, instead of taking a single analytic path for a given dataset, researchers perform and report different analyses for their data and explain how conclusions are affected by the analytical choices (Moors & Hesselmann, 2019). A recent study on unconscious processing used simulations and empirical testing to show that such choices - for example, participant exclusion thresholds and the statistical parameters - can indeed heavily affect the results (Stein et al., 2024). Applying multiverse analysis can help mitigate this problem. Taking the example of participant exclusion, this could involve comparing the results under different scenarios. For example, (a) analyzing the data when including all participants; (b) when excluding participants based on above chance-level performance in the objective awareness measure; (c) when excluding only aware or both insensitive and aware participants using the Bayesian awareness categorization technique (Leganes-Fonteneau et al, 2018); or (d) when excluding participants who report being conscious on many trials in the subjective awareness measure.

Although multiverse analyzes have recently become more frequent (e.g., Steegen et al., 2016; Vadillo, Malejka, & Shanks, 2023), they are still not the norm in the field of cognitive science, including unconscious processing (but see Moors & Hesselman, 2019). The main disadvantage of the multiverse approach is the large number of analyses; this requires more effort, and increases the size of the results section of a single paper.

8. Directly contrasting the processing and objective awareness measures

According to some (Reingold & Merikle, 1988; Meyen et al., 2022), claiming the existence of unconscious processing cannot rest solely on (1) demonstrating a null effect in the objective awareness measure together with (2) an above-null effect in the processing measures, especially given the challenge of establishing a true null effect in (1). According to this view, one should also provide evidence for *a quantitative difference between these two effects*: a larger effect in the processing measure than in the awareness measure. To be able to compare the two measures, however, one needs to place them on the same measurement scale (e.g., since the objective awareness measure is typically binary, one would need to also binarize the processing measures, by, for example, using classification on the reaction time data).

Several studies that followed this logic and directly compared the processing and objective awareness measures failed to find a difference between the two (Meyen et al., 2023; Schnepf et al., 2022, Zerweck et al., 2021; another study found inconclusive results; Huang et al., 2023). This failure to find a larger effect with the processing measure than with the objective awareness measure may be interpreted in one of two ways. One is that the processing effects are not genuinely driven by unconscious processing, but instead are subserved by conscious processing. Alternatively, it might be that the data transformation needed for comparing the two measures reduces the sensitivity of the processing measure, making it harder to find an effect even if it exists. Further research is needed to arbitrate between these two interpretations. Most promising would be to design an objective awareness measure that is as sensitive as the processing measure, as generally speaking, more sensitive awareness measures are needed for a better estimation of conscious processing. However, such a design is yet to be formulated.

	5. Using the liminal or subliminal approach	
	Liminal	Subliminal
Advantages	Can assess the extent to which processes are modulated by conscious perception	Might help avoid trial selection
Disdvantages	Trial selection is necessary, possible priming of the visibility of suppressed stimuli	Cannot readily compare unconscious and conscious processing, possible underestimation of the unconscious effect as the stimuli are more degraded

	6. Dealing with participants having too few trials due to trial exclusion	
	Minimizing trial exclusions at the experimental design stage	Exclude many trials and use appropriate statistical methods to address the exclusion
Advantages	Minimization of data loss	Reduced stimulus suppression, can handle imbalance in the data
Disdvantages	Hard to achieve, might introduce greater physical differences with clearly visible stimuli	Risks of excluding trials (see Fahrenford et al., 2024; Schmidt, 2015), including regression to the mean on the trial level

	7. Performing multiverse analyses
Advantages	Estimates robustness of effects, comparison of results across different definitions of unawareness hereby increasing generalizability
Disdvantages	Requires more effort, and increases the size of the results section of a single paper.

	8. Directly contrasting the processing and objective awareness measures
Advantages	Does not rely on demonstrating a null effect in the objective awareness measure.
Disdvantages	Requires having both measures on the same scale, which typically entails a loss of sensitivity for the processing measure.

Figure 4. A summary of outstanding issues 5-8.

Table 1. Glossary.

Unconscious processing	Processing of a stimulus or a stimulus feature that is not associated with a conscious experience of that stimulus or feature. Claims of unconscious processing typically rely on establishing stimulus processing using a *processing measure* (i.e., behavioral or neural), while providing evidence that the stimulus was not consciously perceived using one or several *awareness measures* (*see below*).
Critical stimulus	The stimulus whose unconscious processing is studied (e.g. the prime in a masked priming experiment")
Processing measure	A measure of performance or neural activity that serves as evidence for the processing of the critical stimulus or a feature of that stimulus.
Direct measure	A measure obtained when participants are explicitly instructed to perform a task on the critical stimulus (or feature) (Reingold & Merikle, 1988). Direct measures can be further subdivided into either subjective or

	objective (Cheesman & Merikle, 1984; Seth et al., 2008). Direct and indirect measures (see below) were originally defined in the context of priming experiments, where an unconsciously processed prime preceded a consciously processed target. In that context, direct measures refer to tasks that pertain to subjectively perceiving the critical stimulus or being able to perform some judgment on it.
Indirect measure	A processing measure (see above) wherein participants are asked to perform a task on a stimulus other than the critical one (Reingold & Merikle, 1988). Performance on that task is used to infer the processing of the critical stimulus. In the context of masked priming experiments, the indirect measure is focused on the behavioral response to a consciously processed target that is either congruent or incongruent with a preceding prime.
Subjective awareness measure	An introspective report of one's perception of the critical stimulus (Ramsøy & Overgaard, 2004; Zehetleitner & Rausch, 2013). Claims of unconscious processing generally rely on participants' reports of not consciously perceiving the stimulus.
Objective measure	A response about the critical stimulus, pertaining either to its existence (i.e., detection) or to its features (i.e., discrimination). When used as an awareness measure, claims of unconscious processing generally rely on providing evidence of *null sensitivity* in this measure (*see below*), alongside another evidence that the stimulus has been processed. The critical difference between subjective and the objective measures is that while the former pertain to reports of one's experience (and can therefore not be externally verified, unless using a forced choice version), the latter pertain to the critical stimulus itself, and can accordingly be deemed correct or incorrect.
Null sensitivity	Demonstration that objective performance is at chance-level. Null sensitivity has often been assessed using statistical tests on individual data to select participants for further analyses (but see the *misclassification* issue below) or at the group level by performing a statistical test on the average performance of each participant.
Contamination by conscious processing	A confound that occurs when the results of the processing measure can be explained by residual awareness of the critical stimulus, which was not captured by both the objective and subjective measures (e.g., due to insufficient sensitivity).
Individual calibration	A procedure whereby the presentation parameters (e.g., intensity or duration of the critical stimulus or mask) are adapted according to each participant's individual perceptual threshold in order to maximize the signal intensity (e.g., Lamy, Salti, & Bar-Haim, 2009; Michel, 2023), while maintaining unawareness.
Subjective	A participant's criterion for reporting subjective perception might lead to

criterion problem	an overestimation or an underestimation of the content of their perception. This might stem from differences in instructions, arousal, motivation, etc. Such a criterion problem can overestimate unconscious processing, if participants report faint perceptions as unconscious.
Misclassification of awareness	A statistical phenomenon where participants or trials are misclassified as “unaware” of the critical stimulus due to measurement error in the awareness measure. If awareness of the critical stimulus or feature is positively correlated with the processing measure, then the accidental inclusion of participants who are truly aware but misclassified as unaware could result in (spurious) evidence of unconscious processing (e.g., due to regression to the mean; Shanks, 2017).

BOX 1 - General Recommendations

The recommendations below apply to many different fields of scientific research, yet we hold them especially important for studies of unconscious processing, as we explain below.

I. Ensure the sample size is sufficient to assure adequate power or evidence to test your hypotheses

Having an adequate sample size is crucial for empirical testing (Fraley & Vazire, 2014; MacCallum et al., 1999) and is especially important when it comes to unconscious processing research, for two reasons. First, as we explain above, strong conclusions are often based on null results (e.g., participants’ performance being at chance-level in the objective awareness measure). Thus, it is crucial to make sure the sample is sufficiently large to detect an effect, if it exists. Second, unconscious effects are sometimes small (Gambarota et al., 2022; Hedger et al., 2016; Van den Bussche et al., 2009), and not always replicable (Biderman & Mudrik, 2018; Moors et al., 2016; Moors & Hesselmann, 2018; Yu, 2023). Thus, studies using frequentist statistics should be powered for the smallest effect one does not want to miss out on (Dienes, 2021). Sample size can be pre-determined in several ways (Lakens, 2022), relying on power analyses. An alternative is to adopt open-ended sequential designs (Schönbrodt et al., 2017), where data collection is stopped once there is enough information to reach a conclusion either for or against unconscious processing, using Bayesian statistics with informed priors (Dienes, 2015).

II. Test your experimental setup systematically

Testing one's setup is very important when studying unconscious processing because the typical paradigms usually require highly accurate presentation timings, given the often very short presentation times (e.g., 10 or 20 ms) or stimulus onset asynchronies. We accordingly suggest that researchers conduct timing tests, for example using an oscilloscope or diode measurement, to test the true presentation duration and the potential on-screen aftereffect of a stimulus. These results should ideally be reported (e.g., in experimental materials) to ensure that discrepant results are not simply explained by unplanned variations in stimulus presentation (Lepauvre et al., 2024).

III. Instruct participants clearly and consistently, and document the instructions

Specifically, for the awareness measures to be interpreted accurately, researchers should take great care in the way that they instruct participants about the experiment. For the subjective awareness measure, this means that it should be clearly defined what stimulus the task refers to (i.e., should awareness be probed for the entire stimulus or a feature of interest?) and that participants should understand the labels that go with each level of the scale (and remember them, in case they are not presented in full on each trial). It should be clear to them that the lowest level means having no conscious perception of the stimulus at all, in case the PAS (or any other visibility rating) is used, or having no confidence at all in one's answer (e.g., "I could have just flipped a coin"), in case a confidence level is used. In the objective awareness measure, it is important to ensure that participants understand that they are being shown a stimulus which is very hard to detect, and in some cases might not even appear, and that it is expected that they do not always consciously perceive it. In addition, the experimenter can explain the importance of trying to answer the question as accurately as possible, even when they did not perceive the stimulus (e.g., by relying on their gut feeling).

IV. Pre-register analyses and share data publicly

Since unconscious processing is generally characterized by small effects (Gambarota et al., 2022; Hedger et al., 2016; Van den Bussche et al., 2009), and replication failures abound, it is particularly important to reduce both false positives and false negatives by protecting against post-hoc decisions that might affect the results (Murayama, Pekrun, & Fiedler, 2014). A more specific problem to the field of unconscious processing is its inherent asymmetry, as studies typically do not distinguish between competing hypotheses, but between a null hypothesis (i.e., no unconscious processing) and the alternative (i.e., unconscious processing). Consequently, such problems as publication bias and selective reporting of successful analyses, over time, lead to the accumulation of false positives. This spuriously increases the corpus of evidence for the scope (rather than limits) of unconscious processing.

For this reason, we recommend that researchers pre-register their analysis plans to reduce degrees of freedom, and share their data publicly in open repositories (or opt for a registered report, where the journal accepts the paper based on the pre-registered methods and analysis plan; Chambers & Tzavella, 2022). Open data - including trial-by-trial information, and not only aggregated data - will also facilitate the pooling of results across studies to precisely quantify the degree of unconscious processing on larger datasets. This recommendation is not meant to discourage exploratory studies; we fully acknowledge the importance of exploration and discovery in science (Coffman & Niederle, 2015). However, we recommend that exploratory experiments be followed by a preregistered replication, or, if this is not a viable option due to some practical constraints, that the results of exploratory analyses are flagged as such and interpreted accordingly.

V. Test for the absence of an effect

As opposed to most empirical tests, where researchers are interested in finding a positive result, the goal of the objective awareness measure is to demonstrate that participants do *not* perform above chance-level (i.e., to confirm the null hypothesis). Classical null hypothesis significance tests are less appropriate in this case because a non-significant result affords two different explanations: (1) the null hypothesis is true, and participants indeed do not consciously perceive the stimuli; or (2) the null hypothesis is false but there is insufficient evidence to reject it. Instead, we recommend that researchers directly test for evidence in favor of the null hypothesis, using either Bayes factors (Dienes, 2016; 2021) or equivalence tests (Kruschke, 2010; Lakens, 2017). Equivalence tests or Bayesian inference by intervals (Kruschke, 2018) require researchers to explicitly state what is the maximum effect size that they consider to be small enough to make a difference; Bayes factors require indicating the size of effect expected. Note that Bayesian statistics might also be helpful to assess the effect of the processing measure (both in the case of a negative and a positive result).

Some further considerations should be taken into account. First, choosing priors for Bayes factors should be done with care (for guidance, see Dienes, 2021; Rouder et al., 2012; Rouder & Morey, 2012). Second, when choosing the bounds for equivalence tests, it is important to keep in mind that when bounds are restrictive, equivalence tests require many trials, but in the end, the bounds should be determined by the smallest effect that would count as scientifically relevant (Lakens, Scheel, & Isager, 2018).

Conclusion

Striking a fine balance between avoiding awareness and allowing processing, and obtaining reliable evidence for both, presents a perennial challenge to the field. This paper provides a list of recommendations for performing research on unconscious processing, pertaining to the design, analysis, and reporting of future studies. It also highlights some of the inherent difficulties in the field; it is evident from this project, and the discussions we held, that current methods are somewhat limited in their ability to assess the extent of conscious vs. unconscious processing, and to manipulate consciousness in a way that both prevents conscious perception of the critical stimulus and preserves a signal strong enough to elicit measurable effects.

Alongside a better understanding of existing methods and measures, innovation is needed; new approaches could potentially break the ongoing cycle of publishing findings that apparently push the limits of unconscious processing, only for them to be pushed back by methodological criticism (e.g., Schmidt, 2015; Shanks, 2017; Stein, 2019). We envision three directions which these novel approaches could follow. First, we should develop new methods to suppress stimuli from awareness while preserving strong signals. Second, there is a need for new measures of awareness or further refinement of existing measures. For example, within subjective awareness measures, it might be beneficial to combine subjective measures with other second-order meta-cognitive measures of performance in some experimental contexts (Jachs et al., 2015; Peters & Lau, 2015; Zehetleitner & Rausch, 2013). Similarly, for the objective awareness measure, it might prove helpful to explore new, non-binary measures, to equate the sensitivity of the objective awareness measure with that of continuous processing measures like reaction times or eye movements (Meyen et al., 2022). Third, information-based approaches that combine machine learning in neuroimaging studies and computational models might prove more sensitive in unraveling unconscious processing than traditional methods (Mei et al., 2022; Mei and Soto, 2024; Sandberg et al., 2022). Developing novel imaging-based processing measures may require substantial effort, yet, if successful, they might overcome some of the above-mentioned limitations (e.g., multi-task costs, confusion between multiple consecutive responses, or order effects).

Until such innovations are introduced, it seems that open questions on the best way to investigate unconscious processing are inevitable. Nevertheless, we believe that following the guidelines we proposed – and refining them in light of future empirical and theoretical work – will improve the reliability and replicability of research on unconscious processing.

Acknowledgements

We thank all researchers who have participated in this process and contributed to the formation of the results. We especially thank Megan Peters and Matthias Michel, who participated in the meetings and in a substantial part of the writing process, but decided to withdraw as authors. L.M. and M.S. are funded by the European Union (ERC, IndUncon, 101077144). N.F. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 803122). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Supplementary Materials

The supplementary materials include two sections. In the first we describe the process by which the recommendations were established; in the second we discuss current methods for addressing the potential threat of misclassification of stimulus visibility due to measurement error.

1. The process by which the recommendations were established

The process of assembling the initial list of recommendations consisted of three polls and three online meetings. These were followed by extensive online discussions while iterating over the manuscript, where the participants provided comments and suggested changes, some of which were extensive. Overall, four separate iterations took place, and the final version was agreed upon by all co-authors. The two polls were sent via email and publicized on social media, and were run on Qualtrics.

1.1. First poll: online distribution via social media and personal emails (N=47)

The purpose of the first poll was to collect researchers' own methodological preferences, what they hold to be the preferred methods in the field, and to understand which issues were most important for us to discuss in this project (See **Figure S1** for the full results). It received 47 responses from researchers in the field from various stages of their career (25 tenured, 12 post-doc, 9 graduate students); most of them came from the disciplines of psychology (30) or neuroscience (15). With regard to the methods, in this initial poll the researchers mostly preferred using both objective and subjective awareness measures as the best method to measure awareness (24; 51%), and using a two alternative forced choice task as the objective awareness measure (21; 45%), specifically targeting the feature of interest (29; 62%); PAS was the most voted on option for the subjective awareness measure (20; 43%). When asked to identify outstanding issues in the way we investigate unconscious processing, they indicated issues pertaining to the measures (36% of the issues indicated; e.g., partial awareness; how to make sure participants were indeed unaware of the stimulus or feature, etc.), methods (32%; e.g., how ecologically valid experiments in the field are; subliminal vs. liminal presentation of the prime, etc.), theory (16%; e.g., inconsistent definitions of the phenomena being studied, etc.) and analysis (16%; e.g., having a measure of evidence of no effect, reducing researcher degrees of freedom, etc.). Researchers were also asked how much they thought recent criticisms of the field (e.g., Meyen et al., 2022; Shanks, 2017) should be addressed in the recommendations, and those considered by the group as important were included in the manuscript.

1.2 Second poll: online distribution via social media and personal emails (N=37)

For the second poll, we compiled a list of methodological issues based on the results of the first poll. For each, we either suggested several alternatives, which the researchers had to choose from, or simply made a recommendation and asked them whether they agreed with it. Each option or recommendation was accompanied by a rationale, and a list of advantages and disadvantages. The second poll received 37 responses (14 tenured, 10 postdocs, 13 graduate students; See **Figure S2** for the full results). Some recommendations emerged, like measuring awareness both subjectively and objectively, using an alternative forced choice discrimination for the objective awareness measure, or focusing the objective awareness measure on the feature of interest. They also highlighted many topics that turned out to be too complicated to make clear-cut decisions about at the time. For instance, when asked whether visible trials should be included in the objective awareness test, 8 voted for the option of including such trials, 7 voted for not including them, and the others refrained from answering; The question of whether or not to exclude participants in different circumstances was also deemed controversial.

1.3 Meetings 1 and 2 (N=32 researchers who attended at least one meeting)

Based on these insights, we formed a new list of recommendations (see below). We then conducted two Zoom meetings with the researchers (N=29 for the first meeting, N=27 for the second meeting) who responded to at least one of the polls to discuss each recommendation in detail and subsequently vote on them. Researchers were sent the list of recommendations in advance of the meeting in order to prepare. During these meetings, each recommendation was presented with its rationale, then a discussion was held, and then it was voted on via Socrative (https://www.socrative.com/).

1.3.1 The original list of recommendations that were voted on during the first two meetings

1. Awareness Measures - Use both a subjective and an objective measure
2. When should the objective test be administered?
   Option 1: take the objective test during the task.
   Option 2: take the objective test in a post-test.
3. Simultaneous objective and subjective measures?
   Option 1: The objective and the subjective measures are obtained via two separate questions.
   Option 2: Both measures are obtained using the subjective-objective measure of

awareness (SOMA; Gelbard-Sagiv et al., 2016)
Option 3: More empirical data should be collected to decide between 1 and 2

4. Subjective measure - PAS, administered on a trial-by-trial basis.
5. Objective measure - An alternative forced choice discrimination (rather than detection), focused on the main feature of interest.
6. Trial exclusion - Should PAS level-2 trials ("brief glimpse") be included as 'unconscious', and under what conditions?
7. Should the test-retest reliability with and without visible primes be tested experimentally?
8. Including visible trials in the objective test
9. Number of trials in objective test - 200 included trials as a lower bound
10. Number of trials in the main task - at least 30 included trials per experimental cell
11. What to do if it is not possible to have enough trials for all subjects?
12. Should sample size be pre-specified (e.g., based on a meaningful effect size or with sequential Bayesian analysis)?
13. How should researchers deal with the problem of excluding subjects?
14. Analysis of all subjects who show above chance performance in the awareness test should also be reported.
15. Is it required to demonstrate an effect over more than one method?
16. Demonstrating the absence of an effect - When assessing the results of the objective measure (or the effect of interest), researchers should not use classical null hypothesis significance tests. Instead, we recommend that they report direct evidence in favor of the null hypothesis.
17. How can researchers reduce degrees of freedom in analyses? - Researchers should pre-register their analyses and share data publicly.

1.4 Third poll - online distribution between co-authors, and Third meeting (N=14)

Following discussions during the two meetings and feedback we received thereafter, we sent a third poll (using Google forms) asking researchers to raise additional topics they thought should be discussed in the final publication. A third meeting was held to discuss these topics, which were presented by the researchers who raised them.

After the third meeting we wrote the first draft of the manuscript. It consisted of the recommendations that reached consensus and outstanding issues that the group had not reached an agreement upon. Recommendations, outstanding issues and their rationale were refined one last time to reflect remaining concerns and objections from the researchers.

2. How to account for regression to the mean at the analysis stage?

Participant exclusion can bias the conclusions of unconscious processing studies. Specifically, when measurement noise causes the misclassification (and thereby: post-hoc exclusion) of participants or trials as “unaware”, this can lead to regression to the mean in the processing measure of the included data. Several solutions have been suggested to address this problem. Given that these solutions are fairly new, we still lack information to arbitrate between them. Instead, we suggest a non-exhaustive, non-mutually exclusive list of options for researchers to explore. These approaches can be divided in two groups, depending on whether they rely on improving participant (or trial) classification, or whether they are aimed at showing that conclusions of unconscious processing cannot be fully explained by regression to the mean. In the former approach, researchers can identify “unconscious participants” using the Bayesian awareness categorization technique (Leganes-Fonteneau et al, 2018), in which a Bayesian procedure is used to classify individual participants as aware, unaware, or insensitive. Drawbacks of this option are that the chosen prior can substantially impact the results, and that even when a large number of trials (e.g., 200) are collected per participant, many will be classified as insensitive. Consequently, this approach might be underpowered (Yaron et al., 2023). In the second group of approaches, researchers do not try to improve the categorization of individual participants as conscious or not, but instead test if the observed effect size is larger than a theoretical effect driven by regression to the mean alone. Here, the following methods have been suggested:

1. Apply a statistical correction for regression to the mean, for example based on Kelley’s formula (Rothkirch et al., 2022) or a Bayesian shrinkage approach (Malejka et al., 2021).
2. The Generative Bayesian framework (Goldstein et al., 2022), which involves modeling the data and adjusting the correlation between performances in the processing measure and objective awareness measure based on the estimated measurement error.
3. A nonparametric bootstrapping approach (Yaron et al., 2023), in which surrogate datasets are simulated to obtain a distribution of effects driven by regression to the mean, allowing researchers to test the hypothesis that the observed effect falls within that distribution.

4. An estimation of the measurement error in the subjective awareness measure using modeling, and accounting for that in the analysis (Skora et al., 2023; see also Jurchis & Dienes, 2023, where this was applied for implicit learning).

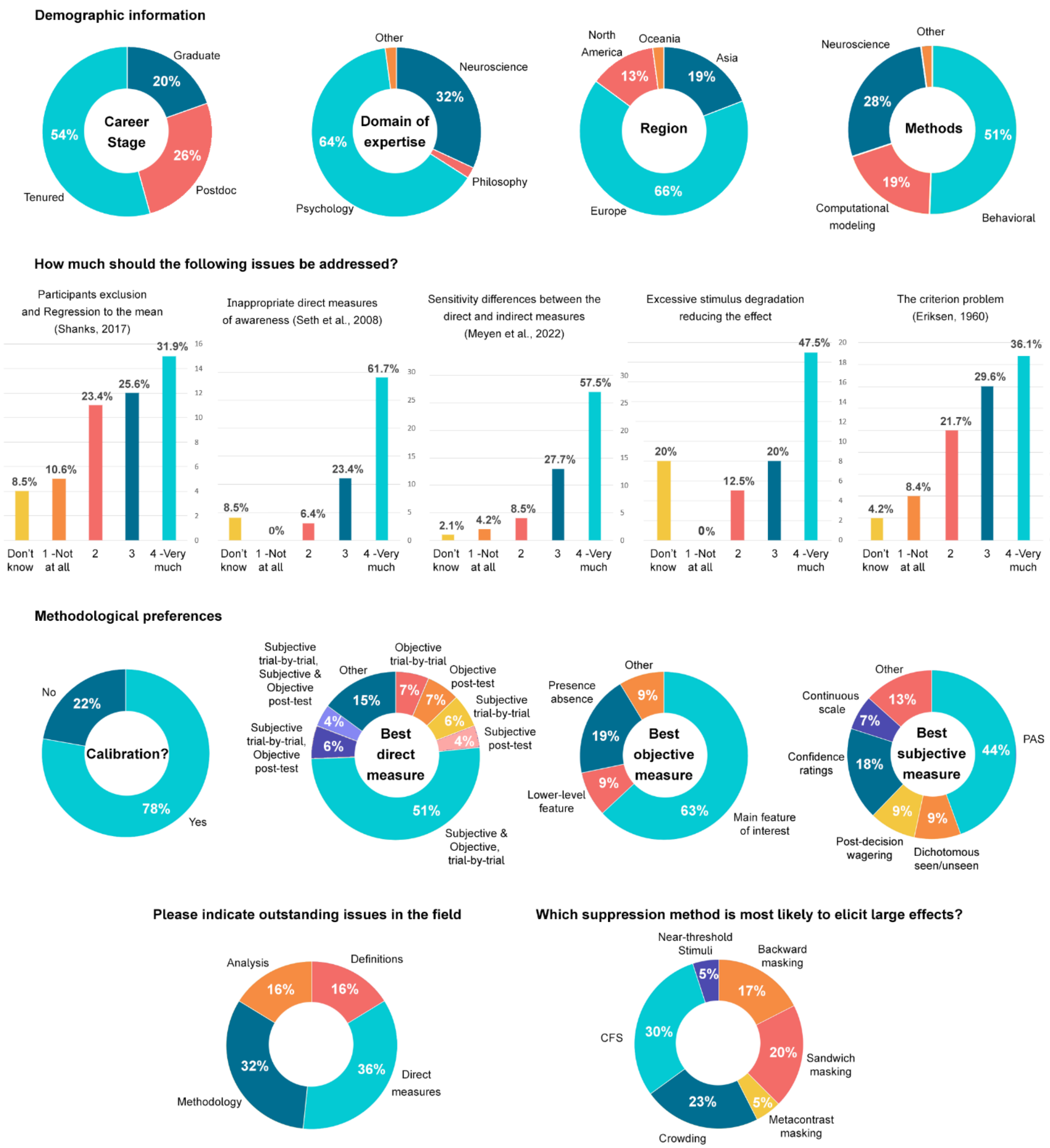

Figure S1. Results from the first online poll (N=47). This poll assessed the methodological preferences of researchers in the field, and its results were the basis for the recommendations that were presented in the second poll.

Figure S2. Results from the second online poll (N=37). Researchers were presented with a preliminary set of recommendations, based on the results of the first poll. The results here were used to form the final set of recommendations that were discussed and voted on during the meetings, as presented in the main manuscript.

References

Alef Ophir, E., Sherman, E., & Lamy, D. (2020). An attentional blink in the absence of spatial attention: a cost of awareness?. *Psychological research, 84(4),* 1039-1055.

Amerio, P., Michel, M., Goerttler, S., Peters, M. A., & Cleeremans, A. (2024). Unconscious Perception of Vernier Offsets. *Open Mind*, *8*, 739-765.

Andersen, L. M., Overgaard, M., & Tong, F. (2019). Visual expectations change subjective experience without changing performance. *Consciousness and Cognition, 71,* 59–69. https://doi.org/10.1016/j.concog.2019.03.007

Andersen, L. M., Pedersen, M. N., Sandberg, K., & Overgaard, M. (2016). Occipital MEG activity in the early time range (< 300 ms) predicts graded changes in perceptual consciousness. *Cerebral cortex, 26(6),* 2677-2688.

Avneon, M., & Lamy, D. (2018). Reexamining unconscious response priming: A liminal-prime paradigm. *Consciousness and cognition, 59,* 87-103.

Avneon, M., & Lamy, D. (2019). Do semantic priming and retrieval of stimulus-response associations depend on conscious perception?. *Consciousness and Cognition, 69,* 36-51.

Axelrod, V., Bar, M., & Rees, G. (2015). Exploring the unconscious using faces. *Trends in cognitive sciences, 19(1),* 35-45.

Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science, 21(2),* 224-233.

Bargh, J. A. (2014). The historical origins of priming as the preparation of behavioral responses: Unconscious carryover and contextual influences of real-world importance. *Social Cognition, 32(Supplement),* 209-224.

Bates, B. T., Dufek, J. S., & Davis, H. P. (1992). The effect of trial size on statistical power. *Medicine & Science in Sports & Exercise, 24(9),* 1059-1065.

Beatty, J., & Moore, A. (2010). Should we aim for consensus?. *Episteme, 7(3),* 198-214.

Bengson, J. J., & Hutchison, K. A. (2007). Variability in response criteria affects estimates of conscious identification and unconscious semantic priming. *Consciousness and Cognition, 16(4),* 785-796.

Biafora, M., & Schmidt, T. (2022). Juggling too many balls at once: Qualitatively different effects when measuring priming and masking in single, dual, and triple tasks. *arXiv preprint* arXiv:2208.13675.

Biderman, D., Shir, Y., & Mudrik, L. (2020). B or 13? Unconscious top-down contextual effects at the categorical but not the lexical level. *Psychological science, 31(6),* 663-677.

Biderman, N., & Mudrik, L. (2018). Evidence for implicit—but not unconscious—processing of object-scene relations. *Psychological science, 29(2),* 266-277.

Breitmeyer, B. G. (2015). Psychophysical “blinding” methods reveal a functional hierarchy of unconscious visual processing. *Consciousness and Cognition, 35,* 234-250.

Brunellière, A., & Hans Frauenfelder, U. (2014). On the locus of grammatical context effects on word recognition. *L’année psychologique, 114(3),* 447-467.

Bruno, N. M., Embon, I., Rivera, M. N. D., Giménez, L., D'Amelio, T. A., Batán, S. T., ... & Andreau, J. M. (2020). Faster might not be better: Pictures may not elicit a stronger unconscious priming effect than words when modulated by semantic similarity. *Consciousness and Cognition, 81,* 102932.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of cognition, 1(1).*

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin, 56(2),* 81.

Chambers, C. D., & Tzavella, L. (2022). The past, present and future of Registered Reports. *Nature human behaviour, 6(1),* 29-42.

Chater, N. (2018). *The mind is flat: The illusion of mental depth and the improvised mind.* Penguin UK.

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & psychophysics, 36(4),* 387-395.

Chng, K. Y., Yap, M. J., & Goh, W. D. (2019). Cross-modal masked repetition and semantic priming in auditory lexical decision. *Psychonomic bulletin & review, 26,* 599-608.

Coffman, L. C., & Niederle, M. (2015). Pre-analysis plans have limited upside, especially where replications are feasible. *Journal of Economic Perspectives, 29(3),* 81-98.

Dai, W., Yang, T., White, B. X., Palmer, R., Sanders, E. K., McDonald, J. A., ... & Albarracín, D. (2023). Priming behavior: A meta-analysis of the effects of behavioral and nonbehavioral primes on overt behavioral outcomes. *Psychological Bulletin, 149(1-2),* 67.

De Gardelle, V., & Mamassian, P. (2014). Does confidence use a common currency across two visual tasks?. *Psychological science, 25(6),* 1286-1288.

Dienes, Z. (2007). Subjective measures of unconscious knowledge. *Progress in brain research, 168,* 49-269.

Dienes, Z. (2015). How Bayesian statistics are needed to determine whether mental states are unconscious.

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology, 72,* 78-89.

Dienes, Z. (2021). How to use and report Bayesian hypothesis tests. *Psychology of Consciousness: Theory, Research, and Practice, 8(1),* 9.

Dienes, Z. (2021). Obtaining Evidence for No Effect. *Collabra: Psychology*, *7*(1).

Dienes, Z. (in press). How to deal with regression to the mean when selecting out conscious trials in order to analyse unconscious trials. *Psychology of Consciousness: Theory, Research, and Practice,* Retrieved from osf.io/preprints/psyarxiv/utrmf

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(5),* 1322.

Dienes, Z., & Seth, A. (2010). Gambling on the unconscious: A comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Consciousness and cognition, 19(2),* 674-681.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. Perspectives on *Psychological science, 1(2),* 95-109.

Dubois, J., & Faivre, N. (2014). Invisible, but how? The depth of unconscious processing as inferred from different suppression techniques. *Frontiers in psychology, 5,* 118154.

Dufek, J. S., Bates, B. T., & Davis, H. P. (1995). The effect of trial size and variability on statistical power. *Medicine and science in sports and exercise, 27(2),* 288-295.

Dulany Jr, D. E. (1962). The place of hypotheses and intentions: an analysis of verbal control in verbal conditioning. *Journal of Personality*, *30*(3).

Dykstra, A. R., Cariani, P. A., & Gutschalk, A. (2017). A roadmap for the study of conscious audition and its neural basis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1714),* 20160103.

Elosegi, P., Mei, N., & Soto, D. (2024). Characterising the role of awareness in ensemble perception. *Journal of Experimental Psychology: Human Perception and Performance, 50(7),* 706–722.

Eriksen, C. W. (1956). Subception: Fact or artifact?. *Psychological Review, 63(1),* 74.

Eriksen, C. W. (1960). Discrimination and learning without awareness: a methodological survey and evaluation. *Psychological review, 67(5),* 279.

Fadardi, J. S., & Bazzaz, M. M. (2011). A Combi-Stroop test for measuring food-related attentional bias. *Experimental and Clinical Psychopharmacology, 19(5),* 371.

Fahrenfort, J. J., Johnson, P. A., Kloosterman, N. A., Stein, T., & van Gaal, S. (2024). Criterion placement threatens the construct validity of neural measures of consciousness. *bioRxiv,* 2024-02.

Faivre, N., & Kouider, S. (2011). Multi-feature objects elicit nonconscious priming despite crowding. *Journal of Vision, 11(3),* 2-2.

Faivre, N., Salomon, R., Vuillaume, L., & Blanke, O. (2016). Numerical priming between touch and vision depends on tactile discrimination. *Perception, 45(1-2),* 114-124.

Filippova, M. G., & Kostina, D. (2020). Dynamics of priming-effect for subliminally presented ambiguous pictures. *Journal of Cognitive Psychology, 32(2),* 199-213.

Fischer, R., Kiesel, A., Kunde, W., & Schubert, T. (2011). Selective impairment of masked priming in dual-task performance. *Quarterly Journal of Experimental Psychology, 64(3),* 572-595.

Fleming, S. M. (2017). HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of consciousness, 2017(1),* nix007.

Fraley, R. C., & Vazire, S. (2014). The N-pact factor: Evaluating the quality of empirical journals with respect to sample size and statistical power. *PloS one, 9(10),* e109019.

Francken, J. C., Beerendonk, L., Molenaar, D., Fahrenfort, J. J., Kiverstein, J. D., Seth, A. K., & Van Gaal, S. (2022). An academic survey on theoretical foundations, common assumptions and the current state of consciousness science. *Neuroscience of consciousness, 2022(1),* niac011.

Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *Journal of Neuroscience, 34(5),* 1738-1747.

Gambarota, F., Tsuchiya, N., Pastore, M., Di Polito, N., & Sessa, P. (2022). Unconscious Visual Working Memory: A critical review and Bayesian meta-analysis. *Neuroscience & Biobehavioral Reviews, 136,* 104618.

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are “implicit” attitudes unconscious?. *Consciousness and cognition, 15(3),* 485-499.

Gayet, S., Guggenmos, M., Christophel, T. B., Haynes, J. D., Paffen, C. L., Sterzer, P., & Van der Stigchel, S. (2020). No evidence for mnemonic modulation of interocularly suppressed visual input. *NeuroImage, 215,* 116801.

Gayet, S., Van der Stigchel, S., & Paffen, C. L. (2014). Breaking continuous flash suppression: Competing for consciousness on the pre-semantic battlefield. *Frontiers in psychology, 5,* 86471.

Gayet, S., Van der Stigchel, S., & Paffen, C. L. (2014). Seeing is believing: Utilization of subliminal symbols requires a visible relevant context. *Attention, Perception, & Psychophysics, 76,* 489-507

Gelbard-Sagiv, H., Faivre, N., Mudrik, L., & Koch, C. (2016). Low-level awareness accompanies “unconscious” high-level processing during continuous flash suppression. *Journal of vision, 16(1),* 3-3.

Giménez-Fernández, T., Luque, D., Shanks, D. R., & Vadillo, M. A. (2023). Rethinking Attentional Habits. *Current Directions in Psychological Science*, *32*(6), 494-500.

Goldstein, A., Sklar, A. Y., & Siegelman, N. (2022). Accurately measuring nonconscious processing using a generative Bayesian framework. *Psychology of Consciousness: Theory, Research, and Practice, 9(4),* 336.

Graham, S., & Lowery, B. S. (2004). Priming unconscious racial stereotypes about adolescent offenders. *Law and human behavior, 28,* 483-504.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969-2012). New York: Wiley.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review, 102(1),* 4.

Guggenmos, M. (2021). Measuring metacognitive performance: type 1 performance dependence and test-retest reliability. *Neuroscience of consciousness, 2021(1),* niab040.

Haase, S. J., & Fisk, G. D. (2020). Extending a focused attention paradigm to critically test for unconscious congruency effects. *Visual Cognition, 28(2),* 148-164.

Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science, 8(2),* 195-207.

Hedger, N., Gray, K. L., Garner, M., & Adams, W. J. (2016). Are visual threats prioritized without awareness? A critical review and meta-analysis involving 3 behavioral paradigms and 2696 observers. *Psychological bulletin, 142(9),* 934.

Hesselmann, G. (2013). Dissecting visual awareness with FMRI. *The Neuroscientist, 19(5),* 495-508.

Hesselmann, G., Hebart, M., & Malach, R. (2011). Differential BOLD activity associated with subjective and objective reports during "blindsight" in normal observers. *Journal of Neuroscience, 31(36),* 12936-12944.

Hesselmann, G., & Moors, P. (2015). Definitely maybe: can unconscious processes perform the same functions as conscious processes?. *Frontiers in Psychology, 6,* 145300.

Hirschhorn, R., Kahane, O., Gur-Arie, I., Faivre, N., & Mudrik, L. (2021). Windows of integration hypothesis revisited. *Frontiers in Human Neuroscience, 14,* 617187.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *The Behavioral and Brain Sciences, 9,* 1-66.

Huang, Q., Jiang, R., Mao, X., Shi, J., & Li, A. (2023). Does response inhibition occur unconsciously? A systematic review and Bayesian meta-analysis. *Consciousness and Cognition, 115,* 103570.

Izatt, G., Dubois, J., Faivre, N., & Koch, C. (2014). A direct comparison of unconscious face processing under masking and interocular suppression. *Frontiers in psychology, 5,* 95953.

Jachs, B., Blanco, M. J., Grantham-Hill, S., & Soto, D. (2015). On the independence of visual awareness and metacognition: a signal detection theoretic analysis. *Journal of experimental psychology: human perception and performance, 41(2),* 269.

Jiang, J., Zhang, Q., & van Gaal, S. (2015). Conflict awareness dissociates theta-band neural dynamics of the medial frontal and lateral frontal cortex during trial-by-trial cognitive control. *Neuroimage*, *116*, 102-111.

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the experimental analysis of behavior, 77(2),* 129-146.

Jurchiș, R., & Dienes, Z. (2023). Implicit learning of regularities followed by realistic body movements in virtual reality. *Psychonomic Bulletin & Review, 30(1),* 269-279.

Kiefer, M., Harpaintner, M., Rohr, M., & Wentura, D. (2023). Assessing subjective prime awareness on a trial-by-trial basis interferes with masked semantic priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49(2),* 269.

Kiefer, M., & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *Neuroreport, 11(11),* 2401-2407.

Kim, C. Y., & Blake, R. (2005). Measuring visual awareness. *Trends in Cognitive Sciences, 8(9),* 381-388.

Kimchi, R., Devyatko, D., & Sabary, S. (2018). Can perceptual grouping unfold in the absence of awareness? Comparing grouping during continuous flash suppression and sandwich masking. *Consciousness and Cognition, 60,* 37-51.

Kimchi, R. Devyatko, D., & Sabary, S. (2023). Perceptual organization and visual awareness: the case of amodal completion. *Front. Psychol.* 14:1201681.

Kouider, S., De Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in cognitive sciences, 14(7),* 301-307.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1481),* 857-875.

Kouider, S., & Dupoux, E. (2001). A functional disconnection between spoken and visual word recognition: Evidence from unconscious priming. *Cognition, 82(1),* B35-B49.

Kouider, S., & Dupoux, E. (2004). Partial awareness creates the “illusion” of subliminal semantic priming. *Psychological science, 15(2),* 75-81.

Kristensen, S. B., Sandberg, K., & Bibby, B. M. (2020). Regression methods for metacognitive sensitivity. *Journal of Mathematical Psychology, 94,* 102297.

Kruschke, J. K. (2010). Bayesian data analysis. Wiley Interdisciplinary Reviews: *Cognitive Science, 1(5),* 658-676.

Kruschke, J. K. (2018). Rejecting or accepting parameter values in Bayesian estimation. *Advances in methods and practices in psychological science*, *1*(2), 270-280.

Kuhn, G., & Dienes, Z. (2006). Differences in the types of musical regularity learnt in incidental-and intentional-learning conditions. *Quarterly Journal of Experimental Psychology, 59(10),* 1725-1744.

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social psychological and personality science, 8(4),* 355-362.

Lakens, D. (2022). Sample size justification. *Collabra: psychology, 8(1),* 33267.

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in methods and practices in psychological science, 1(2),* 259-269.

Lamy, D., Alon, L., Carmel, T., & Shalev, N. (2015). The role of conscious perception in attentional capture and object-file updating. *Psychological science, 26(1),* 48-57.

Lamy, D., Ophir, E. A., & Avneon, M. (2019). Studying the benefits and costs of conscious perception with the liminal-prime paradigm. In *Transitions between consciousness and unconsciousness* (pp. 118-147). Routledge.

Lamy, D., Salti, M., & Bar-Haim, Y. (2009). Neural correlates of subjective awareness and unconscious processing: an ERP study. *Journal of cognitive neuroscience, 21(7),* 1435-1446.

Lanfranco, R. C., Rabagliati, H., & Carmel, D. (2023). Assessing the influence of emotional expressions on perceptual sensitivity to faces overcoming interocular suppression. *Emotion.*

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences, 103(49),* 18763-18768.

Leganes-Fonteneau, M., Scott, R., & Duka, T. (2018). Attentional responses to stimuli associated with a reward can occur in the absence of knowledge of their predictive values. *Behavioural Brain Research, 341*, 26-36.

Leganes-Fonteneau, M., Scott, R., Duka, T., & Dienes, Z. (2021). Avoiding pitfalls: Bayes factors can be a reliable tool for post hoc data selection in implicit learning. *Psychonomic Bulletin & Review,* 1-12.

Lepauvre, A., Melloni, L., Hirschhorn, R., Mudrik, L., & Bendtz, K. (2024). A standardized framework to test event-based experiments.

Lin, Z., & Murray, S. O. (2014). Priming of awareness or how not to measure visual awareness. *Journal of vision, 14(1),* 27-27.

Lloyd, D. A., Abrahamyan, A., & Harris, J. A. (2013). Brain-stimulation induced blindsight: unconscious vision or response bias?. *PloS one, 8(12),* e82828.

Lourenco, S. F., Ayzenberg, V., & Lyu, J. (2016). A general magnitude system in human adults: Evidence from a subliminal priming paradigm. *Cortex, 81,* 93-103.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods, 4(1),* 84.

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. Visual attention, 8(01).

Malejka, S., Vadillo, M. A., Dienes, Z., & Shanks, D. R. (2021). Correlation analysis to investigate unconscious mental processes: A critical appraisal and mini-tutorial. *Cognition, 212,* 104667.

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and cognition, 21(1),* 422-430.

Marcel, A. J. (1974, July). Perception with and without awareness. In *meeting of the Experimental Psychology Society, Stirling, Scotland*.

Maresch, J., Mudrik, L., & Donchin, O. (2021). Measures of explicit and implicit in motor learning: what we know and what we don't. *Neuroscience & Biobehavioral Reviews, 128,* 558-568.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: prestimulus α phase predicts visual awareness. *Journal of Neuroscience*, *29*(9), 2725-2732.

Mei, N., Santana, R., & Soto, D. (2022). Informative neural representations of unseen contents during higher-order processing in human brains and deep artificial networks. *Nature human behaviour, 6(5),* 720-731.

Mei, N., & Soto, D. (2024). Brain representation in conscious and unconscious vision. *bioRxiv* 2024.05.27.596053; doi: https://doi.org/10.1101/2024.05.27.596053

Merikle, P. M. (1984). Toward a definition of awareness. *Bulletin of the Psychonomic Society, 22(5),* 449-450.

Merikle, P. M., & Reingold, E. M. (1990). Recognition and lexical decision without detection: Unconscious perception? *Journal of Experimental Psychology: Human Perception and Performance, 16(3),* 574–583. doi:10.1037/0096-1523.16.3.574

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from Cognitive Psychology. *Cognition, 79(1-2),* 115-134.

Mei, N., & Soto, D. (2024). Brain representation in conscious and unconscious vision. bioRxiv, 2024-05.

Meyen, S., Vadillo, M. A., von Luxburg, U., & Franz, V. H. (2023). No evidence for contextual cueing beyond explicit recognition. *Psychonomic Bulletin & Review,* 1-24.

Meyen, S., Zerweck, I. A., Amado, C., von Luxburg, U., & Franz, V. H. (2022). Advancing research on unconscious priming: When can scientists claim an indirect task advantage?. *Journal of Experimental Psychology: General, 151(1),* 65.

Michel, M. (2023). Calibration in consciousness science. *Erkenntnis, 88(2),* 829-850.

Michel, M. (2023). How (not) to underestimate unconscious perception. *Mind & Language, 38(2),* 413-430.

Micher, N., & Lamy, D. (2023). The role of conscious perception in semantic processing: Testing the action trigger hypothesis. *Consciousness and Cognition, 107,* 103438.

Moors, P., Boelens, D., Van Overwalle, J., & Wagemans, J. (2016, January). Scene integration without awareness: Are object-background relationships extracted for perceptually suppressed scenes. In *Tagung Experimentell Arbeitender Psychologen/Psychologinnen,* Date: 2016/03/21-2016/03/23, Location: Heidelberg, Germany.

Moors, P., Gayet, S., Hedger, N., Stein, T., Sterzer, P., van Ee, R., ... & Hesselmann, G. (2019). Three criteria for evaluating high-level processing in continuous flash suppression. *Trends in Cognitive Sciences, 23(4),* 267-269.

Moors, P., & Hesselmann, G. (2018). A critical reexamination of doing arithmetic nonconsciously. *Psychonomic Bulletin & Review, 25,* 472-481.

Moors, P., & Hesselmann, G. (2019). Unconscious arithmetic: Assessing the robustness of the results reported by Karpinski, Briggs, and Yale (2018). *Consciousness and cognition, 68,* 97-106.

Moran, T., Nudler, Y., & Bar-Anan, Y. (2023). Evaluative conditioning: Past, present, and future. *Annual Review of Psychology, 74,* 245-269.

Mudrik, L., & Deouell, L. Y. (2022). Neuroscientific evidence for processing without awareness. *Annual review of neuroscience, 45,* 403-423.

Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in cognitive sciences, 18(9),* 488-496.

Murayama, K., Pekrun, R., & Fiedler, K. (2014). Research practices that can prevent an inflation of false-positive rates. *Personality and Social Psychology Review, 18(2),* 107-118.

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition, 80(3),* 215-229.

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences, 37(1),* 1-19.

Newell, B. R., & Shanks, D. R. (2023). *Open minded: Searching for truth about the unconscious mind.* MIT Press.

Ortells, J. J., Kiefer, M., Castillo, A., Megías, M., & Morillas, A. (2016). The semantic origin of unconscious priming: Behavioral and event-related potential evidence during category congruency priming from strongly and weakly related masked words. *Cognition, 146*, 143-157.

Overgaard, M. (2015). The challenge of measuring consciousness. *Behavioural methods in consciousness research,* 7-19.

Overgaard, M., & Sandberg, K. (2012). Kinds of access: different methods for report reveal different kinds of metacognitive access. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1594),* 1287-1296.

Park, J. L., & Donaldson, D. I. (2016). Investigating the relationship between implicit and explicit memory: Evidence that masked repetition priming speeds the onset of recollection. *NeuroImage, 139,* 8-16.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature neuroscience*, *4*(7), 739-744.

Peirce, C. S., & Jastrow, J. (1884). On small differences in sensation.

Peremen, Z., & Lamy, D. (2014). Do conscious perception and unconscious processing rely on independent mechanisms? A meta-contrast study. *Consciousness and cognition, 24,* 22-32.

Pernet, C., Garrido, M. I., Gramfort, A., Maurits, N., Michel, C. M., Pang, E., ... & Puce, A. (2020). Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nature neuroscience, 23(12),* 1473-1483.

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature neuroscience, 10(2),* 257-261.

Pesciarelli, F., Scorolli, C., & Cacciari, C. (2019). Neural correlates of the implicit processing of grammatical and stereotypical gender violations: A masked and unmasked priming study. *Biological psychology, 146,* 107714.

Peters, M. A., Kentridge, R. W., Phillips, I., & Block, N. (2017). Does unconscious perception really exist? Continuing the ASSC20 debate. *Neuroscience of consciousness, 2017(1),* nix015.

Peters, M. A., & Lau, H. (2015). Human observers have optimal introspective access to perceptual processes even for visually masked stimuli. *Elife, 4,* e09651.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., ... & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology, 37(2),* 127-152.

Pitts, M. A., Martínez, A., & Hillyard, S. A. (2012). Visual processing of contour patterns under conditions of inattentional blindness. *Journal of cognitive neuroscience, 24(2),* 287-303.

Pitts, M. A., Metzler, S., & Hillyard, S. A. (2014). Isolating neural correlates of conscious perception from neural correlates of reporting one's perception. *Frontiers in psychology, 5,* 105150.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., ... & Yarkoni, T. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature reviews neuroscience, 18(2),* 115-126.

Pratte, M. S., & Rouder, J. N. (2009). A task-difficulty artifact in subliminal priming. *Attention, Perception, & Psychophysics, 71(6),* 1276-1283.

Rabagliati, H., Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. *Journal of Experimental Psychology: General, 147(2),* 190.

Rajananda, S., Zhu, J., & Peters, M. A. (2020). Normal observers show no evidence for blindsight in facial emotion perception. *Neuroscience of consciousness, 2020(1),* niaa023.

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the cognitive sciences, 3,* 1-23.

Rausch, M., Müller, H. J., & Zehetleitner, M. (2015). Metacognitive sensitivity of subjective reports of decisional confidence and visual experience. *Consciousness and cognition, 35,* 192-205.

Rausch, M., & Zehetleitner, M. (2016). Visibility is not equivalent to confidence in a low contrast orientation discrimination task. *Frontiers in psychology, 7,* 183184.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & psychophysics, 44(6),* 563-575.

Ro, T., Singhal, N. S., Breitmeyer, B. G., & Garcia, J. O. (2009). Unconscious processing of color and form in metacontrast masking. *Perception & Psychophysics, 71(1),* 95-103.

Rodríguez-Martínez, G. A., & Castillo-Parra, H. (2018). Bistable perception: neural bases and usefulness in psychological research. *International Journal of Psychological Research, 11(2),* 63-76.

Rothkirch, M., & Hesselmann, G. (2017). What we talk about when we talk about unconscious processing–A plea for best practices. *Frontiers in psychology, 8,* 258414.

Rothkirch, M., Shanks, D. R., & Hesselmann, G. (2022). The pervasive problem of post hoc data selection in studies on unconscious processing. *Experimental Psychology.*

Rothkirch, M., Stein, T., Sekutowicz, M., & Sterzer, P. (2012). A direct oculomotor correlate of unconscious visual processing. *Current biology, 22(13),* R514-R515.

Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research, 47(6)*, 877-903.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of mathematical psychology, 56(5),* 356-374.

Rouder, J. N., Lu, J., Sun, D., Speckman, P., Morey, R., & Naveh-Benjamin, M. (2007). Signal detection models with random participant and item effects. *Psychometrika, 72(4),* 621-642.

Sand, A., & Nilsson, M. E. (2017). When perception trumps reality: Perceived, not objective, meaning of primes drives Stroop priming. *Psychological science, 28(3),* 346-355.

Sandberg, K., Bibby, B. M., Timmermans, B., Cleeremans, A., & Overgaard, M. (2011). Measuring consciousness: task accuracy and awareness as sigmoid functions of stimulus duration. *Consciousness and cognition, 20(4),* 1659-1675.

Sandberg, K., Del Pin, S. H., Bibby, B. M., & Overgaard, M. (2014). Evidence of weak conscious experiences in the exclusion task. *Frontiers in psychology*, *5*, 1080.

Sandberg, K., Del Pin, S. H., Overgaard, M., & Bibby, B. M. (2022). A window of subliminal perception. *Behavioural Brain Research, 426,* 113842.

Sandberg, K., & Overgaard, M. (2015). Using the perceptual awareness scale (PAS). *Behavioral methods in consciousness research*, 181-196.

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other?. *Consciousness and cognition, 19(4),* 1069-1078.

Schmidt, F., Haberkamp, A., & Schmidt, T. (2011). Dos and don’ts in response priming research. *Advances in Cognitive Psychology, 7,* 120.

Schmidt, T. (2015). Invisible stimuli, implicit thresholds: Why invisibility judgments cannot be interpreted in isolation. *Advances in Cognitive Psychology, 11(2),* 31.

Schmidt, T., & Biafora, M. (2022). Explaining the gradient: Requirements for theories of visual awareness. *Behavioral & Brain Sciences,* 45.

Schmidt, T., & Biafora, M. (2024). A theory of visibility measures in the dissociation paradigm. *Psychonomic Bulletin & Review*, *31*(1), 65-88.

Schmidt, T., & Vorberg, D. (2006). Criteria for unconscious cognition: Three types of dissociation. *Perception & Psychophysics, 68(3),* 489-504.

Schnepf, I. A., Friedrich, F., Hepting, C., Meyen, S., & Franz, V. H. (2022). Neural mechanisms of response priming do not support veridical unconscious processing. *Consciousness and Cognition, 102,* 103348.

Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological methods, 22(2),* 322.

Scott, R. B., Samaha, J., Chrisley, R., & Dienes, Z. (2018). Prevailing theories of consciousness are challenged by novel cross-modal associations acquired between subliminal stimuli. *Cognition, 175,* 169-185.

Searle, J. R. (1998). Consciousness, explanatory inversion, and cognitive science. In *Consciousness and Emotion in Cognitive Science* (pp. 139-196). Routledge.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends in cognitive sciences, 12(8),* 314-321.

Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review, 24,* 752-775.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and brain sciences, 17(3),* 367-395.

Skora, L. I., Livermore, J. J., Dienes, Z., Seth, A. K., & Scott, R. B. (2023). Feasibility of unconscious instrumental conditioning: A registered replication. *Cortex, 159,* 101-117.

Soto, D., Mäntylä, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology, 21(22),* R912-R913.

Sperandio, I., Bond, N., & Binda, P. (2018). Pupil size as a gateway into conscious interpretation of brightness. *Frontiers in Neurology, 9,* 416051.

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science, 11(5),* 702-712.

Stein, T., Kaiser, D., Fahrenfort, J. J., & Van Gaal, S. (2021). The human visual system differentially represents subjectively and objectively invisible stimuli. *PLoS biology, 19(5),* e3001241.

Stein, T. (2019). The breaking continuous flash suppression paradigm: Review, evaluation, and outlook. Transitions between consciousness and unconsciousness, 1-38.

Stein, T., & Peelen, M. V. (2021). Dissociating conscious and unconscious influences on visual detection effects. *Nature Human Behaviour, 5(5),* 612-624.

Stein, T., Utz, V., & Van Opstal, F. (2020). Unconscious semantic priming from pictures under backward masking and continuous flash suppression. *Consciousness and Cognition, 78,* 102864.

Stein, T., van Gaal, S., & Fahrenfort, J. J. (2024). How (not) to demonstrate unconscious priming: Overcoming issues with post-hoc data selection, low power, and frequentist statistics. *Consciousness and Cognition, 119,* 103669.

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in cognitive sciences*, *13*(7), 310-318.

Sterzer, P., Stein, T., Ludwig, K., Rothkirch, M., & Hesselmann, G. (2014). Neural processing of visual information under interocular suppression: A critical review. *Frontiers in psychology, 5,* 89205.

Szczepanowski, R., Traczyk, J., Wierzchoń, M., & Cleeremans, A. (2013). The perception of visual emotion: comparing different measures of awareness. *Consciousness and cognition, 22(1),* 212-220.

Timmermans, B., & Cleeremans, A. (2015). How can we measure awareness? An overview of current methods. *Behavioural methods in consciousness research, 21,* 21-46.

Trübutschek, D., Marti, S., & Dehaene, S. (2019). Temporal-order information can be maintained in non-conscious working memory. *Scientific reports, 9(1),* 6484.

Trübutschek, D., Marti, S., Ojeda, A., King, J. R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *elife, 6,* e23871.

Trübutschek, D., Marti, S., Ueberschär, H., & Dehaene, S. (2019). Probing the limits of activity-silent non-conscious working memory. *Proceedings of the National Academy of Sciences, 116(28),* 14358-14367.

Tsikandilakis, M., Bali, P., Derrfuss, J., & Chapman, P. (2019). The unconscious mind: From classical theoretical controversy to controversial contemporary research and a practical illustration of the “error of our ways”. *Consciousness and Cognition, 74,* 102771.

Tsikandilakis, M., Bali, P., Derrfuss, J., & Chapman, P. (2020). “I can see you; I can feel it; and vice-versa”: consciousness and its relation to emotional physiology. *Cognition and Emotion*, *34*(3), 498-510.

Tu, S., Liu, C., Zhu, S., Jou, J., Zhou, Y., & Wan, S. (2019). The semantic integration between two subliminally perceived words simultaneously presented at different locations. *Journal of Psycholinguistic Research, 48,* 1087-1110.

Tunney, R. J., & Shanks, D. R. (2003). Subjective measures of awareness and implicit cognition. *Memory & cognition*, *31*(7), 1060-1071.

Vadillo, M. A., Malejka, S., Lee, D. Y., Dienes, Z., & Shanks, D. R. (2022). Raising awareness about measurement error in research on unconscious mental processes. *Psychonomic Bulletin & Review, 29(1),* 21-43.

Vadillo, M. A., Malejka, S., & Shanks, D. (2023). Mapping the reliability multiverse of contextual cuing.

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychological bulletin, 135(3),* 452.

Van den Bussche, E., Vermeiren, A., Desender, K., Gevers, W., Hughes, G., Verguts, T., & Reynvoet, B. (2013). Disentangling conscious and unconscious processing: A subjective trial-based assessment approach. *Frontiers in Human Neuroscience, 7,* 769.

van Gaal, S., De Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers in human neuroscience, 6,* 121.

Waggoner, J., Carline, J. D., & Durning, S. J. (2016). Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. *Academic Medicine, 91(5),* 663-668.

Wang, X., & Forster, K. (2015). Is translation priming asymmetry due to partial awareness of the prime?. *Bilingualism: Language and Cognition, 18(4),* 657-669.

Wierzchoń, M., Anzulewicz, A., Hobot, J., Paulewicz, B., & Sackur, J. (2019). In search of the optimal measure of awareness: Discrete or continuous?. *Consciousness and cognition, 75,* 102798.

Wierzchoń, M., Paulewicz, B., Asanowicz, D., Timmermans, B., & Cleeremans, A. (2014). Different subjective awareness measures demonstrate the influence of visual identification on perceptual awareness ratings. *Consciousness and cognition, 27,* 109-120.

Williams, J. N. (2020). The neuroscience of implicit learning. *Language Learning, 70(S2),* 255-307.

Yang, Y. H., Tien, Y. H., Yang, P. L., & Yeh, S. L. (2017). Role of consciousness in temporal integration of semantic information. *Cognitive, Affective, & Behavioral Neuroscience, 17,* 954-972.

Yaron, I., Faivre, N., Mudrik, L., & Mazor, M. (2023). Individual differences do not mask effects of unconscious processing.

Yaron, I., Melloni, L., Pitts, M., & Mudrik, L. (2022). The ConTraSt database for analysing and comparing empirical studies of consciousness theories. *Nature Human Behaviour, 6(4),* 593-604.

Yaron, I., Zeevi, Y., Korisky, U., Marshall, W., & Mudrik, L. (2023). Progressing, not regressing: A possible solution to the problem of regression to the mean in unconscious processing studies. *Psychonomic Bulletin & Review,* 1-16.

Yu, R. (2023). Unconscious integration: Current evidence for integrative processing under subliminal conditions. *British Journal of Psychology, 114(2),* 430-456.

Zehetleitner, M., & Rausch, M. (2013). Being confident without seeing: what subjective measures of visual consciousness are about. *Attention, Perception, & Psychophysics, 75,* 1406-1426.

Zerweck, I. A., Kao, C. S., Meyen, S., Amado, C., von Eltz, M., Klimm, M., & Franz, V. H. (2021). Number processing outside awareness? Systematically testing sensitivities of direct and indirect measures of consciousness. *Attention, Perception, & Psychophysics, 83(6),* 2510-2529.