The challenge of inferring unconscious mental processes

David R. Shanks
University College London, England

Simone Malejka
University of Cologne, Germany

Miguel A. Vadillo
Universidad Autónoma de Madrid, Spain

Address correspondence to:

David R. Shanks
Division of Psychology and Language Sciences
University College London
26 Bedford Way
London WC1H 0AP
d.shanks@ucl.ac.uk

Acknowledgments: This research was supported by a grant from the United Kingdom Economic and Social Research Council (ES/P009522/1), and MAV was supported by grants 2016-T1/SOC-1395 and 2020-5A/SOC-19723 (Comunidad de Madrid) and PSI2017-85159-P (AEI / FEDER UE). We thank two anonymous reviewers for helpful suggestions on this work. The R code for the models in Table 1 and Figure 1 is available via the Open Science Framework (OSF) at https://osf.io/hm4ta/.

Abstract

Studies of unconscious mental processes often compare a performance measure (e.g., some assessment of perception or memory) with a measure of awareness (e.g., a verbal report or forced-choice response) of the critical cue or contingency taken either concurrently or separately. The resulting patterns of bivariate data across participants lend themselves to several analytic approaches for inferring the existence of unconscious mental processes, but it is rare for researchers to consider the underlying generative processes that might cause these patterns. We show that bivariate data are generally insufficient to discriminate single-process models, with a unitary latent process determining both performance and awareness, from dual-process models, comprising distinct latent processes for performance and awareness. Future research attempting to isolate and investigate unconscious processes will need to employ richer types of data and analyses.

KEY WORDS: awareness; bias; correlation; regression to the mean; unconscious

Research on the nature and scope of unconscious or implicit mental processes has long been a core concern of cognitive psychologists and interest in the topic shows no signs of abating (e.g., Blake, 2021; LeDoux et al., 2020). Reasons for this interest are not hard to find. For example, our perceptual experience is strikingly impoverished (Cohen et al., 2021), raising the question of how much impact events have on us when they are not consciously perceived. Similarly, we appear to be able to acquire complex skills such as language with little awareness of those rules (Chomsky, 1980).

When researchers explore unconscious influences in learning, memory, action, perception, decision making, cognitive control, emotion, and other domains, they commonly employ experimental tasks in which two measurements are taken. One is the primary performance index and the other is a measurement of the participant's awareness of whatever critical stimulus feature or relationship is assumed to drive performance in the primary task. For instance, speed of responding to briefly-presented angry or neutral faces might be measured together with perceptual awareness (e.g., discrimination) of facial expression (Hedger et al., 2016), or learning of an advantageous decision strategy might be related to verbal reports of the strategy (Bechara et al., 1997). Although there are other general approaches for studying unconscious processes – for example, research on the Unconscious Thought Effect and the Implicit Association Test (IAT) take different approaches; see Abadie and Waroquier (2019) and Greenwald and Banaji (2017) – many hundreds of studies conducted over the past several decades can be conceptualised within this framework of bivariate measurement.

Given how common the practice is of collecting bivariate performance/awareness data, one might imagine that there has been rich discussion about the general strategies available to researchers for analysing and interpreting such data. After all, inferring unobservable mental processes from patterns of observable behaviour is one of the most fundamental and challenging problems facing psychological science (Kellen et al., 2021). Yet with rather few exceptions (e.g., Dienes, 2015; Greenwald et al., 1995; Sand & Nilsson, 2016; Stephens et al., 2019), very little attention has been devoted to evaluating the adequacy of different methods for inferring unconscious mental processes from bivariate data or to examining the assumptions on which they rest.

We undertake such an assessment in the current article. By and large we take a relatively abstract viewpoint, describing particular experimental results only when they help to illustrate a general point. The article is not intended as a review of research on unconscious mental processes, which can be found elsewhere (e.g., Greenwald & Banaji, 2017; Hassin, 2013; Hedger et al., 2016; Lovibond & Shanks, 2002; Mertens & Engelhard, 2020; Newell & Shanks, 2014; Nobre et al., 2020; Phillips, 2021; Weiskrantz, 1997). Instead, we critically examine the adequacy of various inferential approaches that are almost taken for granted by many researchers in this field. We show that the inferential challenge is far more profound than is generally acknowledged, and offer suggestions for how future research on this fundamental topic might surmount this challenge.

There has of course been extensive debate about what constitutes an acceptable method for measuring awareness, evaluating the relative merits of verbal reports, visibility ratings, discrimination at the subjective and objective thresholds, confidence reports and so on (Blake, 2021; Cheesman & Merikle, 1984; Dienes & Scott, 2005; Ericsson & Simon, 1984; Merikle & Reingold, 1992; Peirce & Jastrow, 1884; Sandberg et al., 2010; Shanks & St. John, 1994). Here we ignore these deep issues and take it as read that data have been collected employing a suitable awareness measure.

Theoretical and conceptual context

Construct validation and formal models play crucial roles in the enterprise of inferring mental processes (Cronbach & Meehl, 1955; Grahek et al., 2021). The general logic has two aspects. The first is to articulate theoretical constructs, employ psychometric methods to measure them reliably, and establish their validity. The second is to place them within a nomological framework, that is, to formulate a model embodying the constructs and their putative interrelationships, although invariably simplifying the domain in which it is to be applied. This model can be purely statistical (e.g., regression, factor analysis, structural equation modelling) or generative (e.g., a computational model is constructed that generates simulated data which can then be compared against the actual data collected empirically). To the extent that model and data align quantitatively and qualitatively, the model gains some support, especially if it outperforms alternative models, though of course it can never be proven that the model is ‘true’. The more complex the data set, in terms of the number of manipulated independent variables and outcome measures, the harder the task for the modeller, but the more meaningful is a good fit to the data.

From the perspective of implicit or unconscious mental processes, research on the IAT provides an example of the application of this general approach. The theory that the IAT provides a measure of implicit attitudes that are dissociable from explicit attitudes can be tested by collecting data on participants' performance on the IAT and related tests as well as their responses on standard explicit attitude questionnaires. The psychometric properties of these measures are assessed. Analytic techniques such as structural equation modelling can then be employed to test different models, in particular to compare a model that has a single latent attitude that determines responses on both implicit and explicit tests versus one that has distinct latent attitudes, one controlling explicit attitudes and the other implicit ones. Manipulations can be introduced which are intended to have selective influence on one type of attitude or the other. Although there is much controversy about the resulting data patterns and their interpretation (Kurdi et al., 2019; Nosek & Smyth, 2007; Schimmack, 2021), the general methodology is well-understood and relatively uncontroversial. For examples from different domains of unconscious processing, see West et al. (2021) for language development and Kaufman et al. (2010) for implicit learning.

When we look more broadly at the methods researchers use to establish the existence or otherwise of unconscious mental processes, we soon find that the above example is the exception rather than the rule. Instead of complex datasets involving multiple independent variables and outcome measures, psychometrically analysed and validated, and contrasted against the predictions of competing theoretical models, researchers collect very simple bivariate data and approach these data from a straightforward analytic perspective. For example, in one of the most influential studies on unconscious processes, Bechara et al. (1997) obtained data across participants on a measure of performance (selecting advantageous cards in risky choice in the Iowa Gambling Task) and a measure of awareness (responses to probe insight questions), and concluded that advantageous selections were made when participants lacked awareness of the choice payoffs. Similarly, in another hugely influential study, Warrington and Weiskrantz (1968) measured performance on a priming measure (identifying degraded pictures or words) and a conscious memory measure across individuals with amnesia, and reported priming in the absence of detectable conscious memory. As a final example, Dehaene et al. (1998) reported that briefly-presented and masked ('subliminal') words affected semantic judgments even though participants performed at chance on an assessment of awareness of the masked words. Of course these studies were much more sophisticated than this very brief overview suggests, but from the perspective of their inferential logic they all relied on interpretation of patterns observed in bivariate data. Similar data form the basis of inference in numerous other studies published over the past 50 years.

In adopting such a minimalist form of data, these studies draw heavily for their inferential justification on the dissociation logic that had its heyday in the 1980s and 1990s (Dunn & Kirsner, 1988; Shallice, 1988). It was the hope of this research program that mental structure could be revealed by patterns of selective influence. For example, if damage to a particular brain structure led to deficits in performing task A but not task B (or even better, if damage to brain structure 1 led to deficits in performing task A but not task B, while damage to brain structure 2 led to deficits in performing task B but not task A), then functionally distinct mental and neural systems could be identified. Amongst many lines of evidence speaking against this strong form of inference (e.g., Kinder & Shanks, 2003; Munakata, 2001; Plaut, 1995), developments in state-trace analysis (Newell & Dunn, 2008; Stephens et al., 2019; Yeates et al., 2015) showed that dissociation patterns are rarely inconsistent with explanations based on a single latent construct that has monotonic but distinct mappings onto the measured dependent variables. For this reason, as well as the finding that many classic examples in the implicit/explicit domain have proven hard to replicate (e.g., Berry et al., 2014; Berry et al., 2017; Le Pelley et al., 2019; Ward et al., 2020; Zaki et al., 2003), dissociation logic has largely been abandoned in recent years. We return to this issue later.

Bivariate data: The basic problem

What is the problem with using bivariate performance/awareness data as the basis for inferences about unconscious processes? At first glance, the logic seems straightforward. Consider the artificial dataset shown in Figure 1A. Each datapoint represents a single participant, with their mean performance (such as response facilitation) across a set of trials on the y-axis plotted against a mean measure of awareness of a critical stimulus feature or relationship on the x-axis. The scale values are unimportant, but we plot performance on a scale going up to hundreds (as might be appropriate for a millisecond measure based on response times [RTs]) and awareness on a scale that might be relevant for a discriminability measure such as *d*′ from signal detection theory. In both cases we assume that zero reflects the baseline or chance level of that measure. Figure 1A shows a 'holy grail' pattern in which the vast majority of a large sample of participants score above chance on the performance measure while scoring at chance on the awareness measure. Indeed, these data are generated from a model in which performance is truly unrelated to awareness, and awareness is completely absent.

Specifically, for a given participant we assume that performance $P = a \cdot S_P + e_P$ where S_P is a normally distributed variable representing whatever latent mental content or representation determines the individual's level of performance strength, a is a constant scaling factor, and e_P is a random error term (see Table 1). This error represents all potential sources of noise, both intrinsic to the individual such as trial-by-trial variability in attention as well as extrinsic such as measurement error. At the same time, we calculate the participant's level of awareness from the equation $A = b \cdot S_A + e_A$ where S_A represents the latent mental content that determines the individual's measured level of awareness, b is another constant scaling factor, and e_A is another source of random error. These equations may seem unnecessarily complex as the S and e terms in each formula could trivially be replaced by a single variable, but we keep them distinct to emphasize that S represents an unobservable latent construct, the very entity that our analysis is attempting to identify, and that this construct varies across individuals – some perform better and some worse in the experimental task. Table 1 gives details of the means and standard deviations of the normal distributions from which the various parameters are sampled and the values of a and b. None of the points made in this article depend particularly on these values – they are chosen to yield patterns that bear some approximate similarity to typical experimental results. Note that the error terms always have a mean of zero. This model is designated Model 2 to emphasize the fact that it includes 2 latent processes, S_P and S_A.

What exactly do the S parameters represent? They are intended to be very general theoretical constructs that simply reflect whatever unobservable mental state determines a particular measurable behaviour, be it on a performance or an awareness scale. In visual perception, S could represent the amount of information picked up by the senses about an external object. In the case of learning and memory, it could represent the state of acquired/retained knowledge extracted from events, akin to how familiarity is conceptualized in signal detection theory. In skill acquisition it could represent the degree of acquired domain expertise. We simply assume that these latent constructs have normally-distributed values across participants. Although the error terms in the model affect performance and awareness, they do so in a way that is completely independent of the individual's strength of latent knowledge.

It should be clear from these model formulae that the latent entities S_P and S_A are independent. Moreover, since the true mean of S_A is zero at the group level, this model embodies complete independence between performance and awareness as well as a complete absence of awareness and hence licenses quite a strong inference that whatever is controlling performance (that is, S_P) is unrelated to awareness. So the data depicted in Figure 1A would form the basis of a compelling argument for the existence of an unconscious process influencing performance. Of course this is not a deductive inference, and it would be entirely open to an opponent of unconscious processes to formulate a model which generates a close fit to the data without the need to distinguish S_A and S_P. Although this would be challenging, it would not be impossible, especially if different functions are incorporated to map *S* onto the two observable measures (Kellen et al., 2021; Stephens et al., 2019). So what is the problem?

Even casual acquaintance with the literature on unconscious or implicit processes makes it clear that patterns like that shown in Figure 1A are as elusive as unicorns. Indeed the frequency with which advocates of unconscious mental processes employ the alternative methods we describe later suggests that they regard this pattern as hard to demonstrate. We will not undertake a detailed literature review here, but it is safe to say that there exist vanishingly few examples of studies which meet the relevant requirements, namely (1) that performance *P* is robustly above chance, (2) it is beyond dispute that awareness *A* is at chance, and (3) that these first two requirements are regularly and convincingly corroborated in replication studies. The requirement that it is beyond dispute that *A* is at chance could be achieved within a frequentist framework either by a hypothesis test with a small fixed Type-II error rate showing a non-significant result (Neyman & Pearson, 1933) or by equivalence tests showing that *A* is smaller than some minimal theoretically-meaningful level (Lakens et al., 2018), or within a Bayesian framework by strong evidence for the null hypothesis (Sand & Nilsson, 2016).

Some patterns that appear to resemble Figure 1A disappear in replications where it turns out that although awareness is truly at chance, performance is also at chance (e.g., Heycke et al., 2017; Röseler et al., 2021; Stein et al., 2020). Many behaviour priming effects – subtle influences on behaviour assumed to lie outside awareness – have disappeared in replication studies (e.g., Doyen et al., 2012; Klein et al., 2014; O'Donnell et al., 2018; Shanks et al., 2015). Meta-analyses also suggest that some implicit effects are inflated by publication bias, with low performance levels being under-represented (e.g., Lodder et al., 2019; Mertens & Engelhard, 2020; Nobre et al., 2020; Shanks et al., 2015; Vadillo, Hardwicke, et al., 2016).

Putting aside the replicability of performance effects in individual studies, another problem inherent in data patterns like that shown in Figure 1A is that they only license an interpretation in terms of unconscious processes if one is willing to assert that the null hypothesis of the awareness test (awareness = 0) is true. If statistical power is low, because a small sample of participants has responded to a small number of awareness trials, then failing to reject a false null hypothesis is a highly likely outcome (Sand & Nilsson, 2016). This is a methodological rather than theoretical concern that can be ameliorated by testing large samples or combining data in a meta-analysis.

When such measures are taken to boost power, it is common to find that awareness is robustly better than chance. Well-powered replications and meta-analyses reveal that awareness is actually above chance in many experiments supposedly studying unconscious processes, conforming to a pattern more like that shown in Figure 1B. For example, studies following Bechara et al.'s (1997) research on the Iowa Gambling Task found that participants often have appreciably high levels of awareness undetected by the methods employed in the original experiments (Konstantinidis & Shanks, 2014; Maia & McClelland, 2004), while a replication of Dehaene et al.'s (1998) subliminal semantic priming effects found a similar result (Meyen et al., in press). In meta-analyses we have found that awareness is reliably above chance when data are aggregated from contextual cuing (Vadillo, Konstantinidis, et al., 2016) and probability cuing (Vadillo et al., 2020) (see also Jiang et al., 2018) experiments almost all of which individually claim the opposite, and Nobre et al. (2020) found a similar pattern in studies on implicit effects in inattentional blindness.

Occasionally researchers have endeavoured to engineer a pattern like that in Figure 1A by deliberately manipulating key aspects of the task in order to reduce awareness to chance (e.g., Stein et al., 2021; Wildegger et al., 2015). For example, subliminal perception studies (e.g., Beauny et al., 2020; Wildegger et al., 2015) sometimes use a 'staircase' procedure in which the presentation duration of a stimulus is increased or decreased across trials until identification accuracy is at chance in every participant. It is unfortunate that this procedure is not employed more often. Even when it is adopted, the results are often inconclusive (we discuss Wildegger et al.'s results in more detail below).

As might be expected from the combination of both above-chance performance and awareness, the data in Figure 1B were generated from a model (Model 1a in Table 1) with a single latent process simultaneously determining both outcome measures[1]. Specifically, performance is based on $P = a \cdot S + e_P$ and awareness on $A = b \cdot S + e_A$, where S no longer has a subscript because it is common to both outcomes. As before, S varies across participants, reflecting the fact that the level of latent task knowledge is systematically greater in some than in others. The noise parameters e_P and e_A are again independent of S and have been chosen to reflect the fact that awareness scores are often found to be much more dispersed than the idealized pattern in Figure 1A.

It should be clear from these illustrations that falsifying a model that only includes a single latent process is not a trivial task. The first two recommendations in Box 1 would alleviate some of the concerns expressed in this section.

Post hoc *selection*

[1] Single-process models are sometimes inappropriately described as ones in which awareness causes performance Sisk, C. A., Remington, R. W., & Jiang, Y. V. (2020). A spatial bias toward highly rewarded locations is associated with awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*, 669-683. https://doi.org/10.1037/xlm0000749 . Awareness is either a behavioural measure such as a report or a subjective state (or both), but in neither case can it cause the performance measure.

As a reaction to the difficulty of obtaining a precise group-level null awareness effect, researchers often adopt what seems at first sight to be an entirely reasonable analytic approach, namely to eliminate from the analysis all participants who score above some more or less strict cut-off on the awareness assessment. Then the main analysis computes the average score on the primary performance measure for the remaining participants, who, seemingly, must lack awareness. If this average performance score exceeds chance, then it appears that unconscious processes must be playing a role. Figure 1C shows such a pattern where only participants scoring at or below chance on the awareness measure are retained.

For example, Wildegger et al. (2015, Exp. 1b) measured the effect of a briefly-presented prime (an oriented line) on speed of reporting the orientation of a target line. The prime's orientation was either congruent or incongruent with that of the target. In a separate prime awareness test, participants made forced-choice prime identification responses and their overall accuracy was determined. Wildegger et al. found that 5 out of 20 participants performed better than chance in this awareness test and hence eliminated them from the primary analysis, which nevertheless obtained a 6 msec priming effect. Wildegger et al. therefore interpreted their findings as evidence of an unconscious influence of the prime.

Shanks (2017) showed that, counter-intuitively, this analytic method is flawed and can lead to wholly false conclusions about unconscious mental processes. Box 2 describes a statistical problem to illustrate the underlying issue. Readers are invited to reflect on it before continuing.

Is there any problem with the apparently intuitive procedure described in Box 2? Indeed there is. To illustrate this concretely, suppose the true toxicities are sampled from a normal distribution with mean of 50 units and a standard deviation (*SD*) of 10 units, and that your equipment adds measurement error with a mean of zero and *SD* = 5. Your estimate of the toxicity of the 5 least toxic additives will be about 36.4, but the true toxicity of these additives is in fact appreciably higher, mean = 39.1.[2]

[2] Moreover, although less relevant to the present concerns, there is another problem with the estimate: because of measurement error, these 5 additives are not necessarily the least toxic ones. Some additives which should be in the set of least toxic ones will be omitted because of the error with which the true toxicities were measured. The mean toxicity of the truly least toxic foods is about 37.8.

This mis-measurement is a consequence of regression to the mean. Although error in the measurement of toxicity is statistically independent of true toxicity, these errors are not independent of the true toxicity of the least toxic additives. This is simply because random under-estimation of an additive's toxicity increases the chances that that additive is selected for inclusion and, conversely, random over-estimation decreases an additive's chances of inclusion. If one were to order the additives from highest to lowest by measured toxicity, the errors would tend to be positive in the most toxic additives and negative in the least toxic ones. Crucially, if you re-measure the toxicity of the 5 chosen additives, their toxicities will increase. On this second, independent measurement, errors will now be genuinely independent of the true toxicities (as well as being independent of the errors in the first measurement) and so the measured toxicities will regress to the mean.

The relevance to studies of unconscious mental processes is straightforward. It could be the case that the measured awareness (and performance) in the hypothetical participants in Figure 1C is precise and accurate. Maybe the awareness test is exceptionally reliable and only a very small amount of error (e_A) is present. But it could equally be the case that error is nontrivial. In this situation, by selecting participants who fall below some awareness cut-off, the researcher is inevitably (on average) selecting ones for whom true awareness is greater than measured awareness, often by a large amount. Suppose that a participant's true latent knowledge is greater than chance, even by only a small amount (that is, $S > 0$). If the error term e_A for this individual is negative, then measured awareness ($= b \cdot S + e_A$) may be less than or equal to chance (0) and the participant will be included in the analysis. Because participants are selected *post hoc* in this method according to their measured awareness, the researcher has no control over the magnitude of error that is inherent to the measured awareness of the participants retained in the analysis. Thus we cannot straightforwardly know whether these participants are truly unaware (e_A is negligible) or truly aware (e_A is negative).

These two alternatives make divergent predictions regarding an easily-tested outcome, namely test-retest reliability. If e_A is genuinely small and measured awareness in the selected participants is an accurate reflection of true awareness, then the results should look very similar on a second, independent test of awareness (high reliability). However if e_A is negative, we will expect regression to the mean to occur on an independent awareness test.

On such a test the values of e_A will have a mean of zero, causing measured awareness ($= b \cdot S + e_A$) to increase from its negative value and regress towards the mean level of awareness of the entire unselected group. The net effect will be low test-retest reliability. In case any readers believe that awareness tests are usually sufficiently reliable to make any such regression effects negligible, we will later describe a published study in which the reliability of the awareness test was only slightly above zero.

The consequences of regression to the mean are well-known (Campbell & Kenny, 1999; Morton & Torgerson, 2005). The following analogy may be helpful. Imagine that a school administers a test to all its students, perhaps a test of executive skills, and on the basis of their normed test scores a group of low-scoring students is assessed for some remedial intervention. After the intervention, their scores are found to have improved appreciably. Readers will immediately recognise that attributing the improvement to the intervention is unwarranted, because it may alternatively be the result of regression to the mean. Some students, purely by chance, will have underperformed on the initial test. Perhaps they were tired or distracted. Thus their scores, which combine a true score plus error, will regress to the mean on a second test in which the random errors are drawn independently from the error distribution. In the context of research on unconscious processes, the mere possibility of such a state of affairs is sufficient to entirely undermine many studies that have based their conclusions of *post hoc* selection without assessing the reliability of their awareness tests (Shanks, 2017).

It might be assumed that any such regression effect will be small, perhaps too small to matter. We will comment later on the actual level of reliability that is seen in typical tests of awareness, but models can once again be instructive. In fact, as perceptive readers will have noticed, the points in Figure 1C are identical to those falling to the left of $A = 0$ in Figure 1B – they were generated from a model with only a single latent process (Model 1A). So we can ask, how much regression would we see in these datapoints? Would it be enough to change our inference that awareness is not greater than zero in these simulated participants? The purple triangle in Figure 1C shows the true mean awareness (and observed mean performance) in the included datapoints. Not only is it shifted to the right, but it is far beyond $A = 0$! In other words, if we were able to strip away the error term that contributes to each datapoint in the figure to leave just the term $b \cdot S$, we would find it to be much less negative – indeed, most of the resulting points would be positive. In fact the mean true awareness is virtually identical to the whole group's observed mean (Figure 1B). Of course this is an extreme case arising from parameters chosen to make a point. With relatively less error in the awareness measure, the degree of regression would be smaller, though always yielding true mean awareness greater than 0 (because mean *S* is greater than zero).

This simple simulation yields a profoundly counterintuitive conclusion. A subset of (simulated) participants can be selected *post hoc* on the basis that their awareness is not above chance, and the mean performance of these participants will be appreciably above chance, indeed close to the overall group mean, apparently revealing an unconscious mental process. Yet the apparent unawareness in these participants is an illusion, caused directly by the *post hoc* selection method. These participants only appear to be unaware. In truth the latent process (*S*) that determines awareness scores is robustly positive. But it has by chance been combined with strongly negative error scores, and then selected by the researcher. If we were to independently re-measure awareness in these participants, regression to the mean would play its inevitable role and reveal their true awareness to be robustly above chance.

To this point our conceptualisation of bivariate data patterns relevant to unconscious processing has assumed that the measure of awareness is continuous, and this is also reflected in the models in Table 1 where *A* is a continuous measure. Researchers (e.g., Beauny et al., 2020; Biderman et al., 2020) frequently however collect discrete (often binary) awareness reports, for example asking participants whether a masked stimulus was 'seen' or 'unseen', or employ scales with more than two categories such as the four-point Perceptual Awareness Scale (Ramsøy & Overgaard, 2004). These represent special cases of bivariate data in which measures on the awareness scale are discrete and the analysis averages across trials rather than participants. Importantly, as Shanks (2017) and Schmidt (2015) have pointed out, treating invisibility reports as true indices of unconscious processing will often lead to incorrect inferences. Eriksen's (1960) famous critical appraisal of unconscious discrimination published over 60 years ago has been described as a "constant reminder [...] that one must maintain a healthy skepticism about subjective reports of what can be seen and what cannot, meaning that one must redouble the effort to validate 'invisibility' when claiming that people do not see things that nonetheless influence their performance on behavioral tasks involving putatively 'invisible' stimuli" (Blake, 2021, p. 546).

Schmidt's (2015) argument is that interpreting invisibility reports at face value pays no heed to the individual's response bias, and decades of development of signal detection theory have demonstrated that discrimination and bias are distinct entities. Responding 'invisible' may not mean that the internal representation of the stimulus is truly unconscious, it may instead mean that its strength is below the threshold the individual has set for reporting 'visible'. Imagine a hypothetical situation in which a participant is tested twice under otherwise identical conditions, except that a standard manipulation of response bias (such as instructions or payoffs) induces a conservative bias in the first test (very few stimuli are reported as visible) and a liberal bias in the second test (many stimuli reported as visible). Consequently, some events will be reported as invisible in the first test but visible in the second. Would it not be highly implausible to infer that the manipulation affects the participant's actual subjective experience? Much more likely is that the underlying state of experience is the same and that it is the reporting criterion that has changed. This response bias alternative is familiar as an explanation of 'blindsight' (Phillips, 2021).

Shanks' (2017) point is distinct from but complementary to this. Reported awareness depends on both the underlying true representation (S) and random error. Stimuli that evoke 'invisible' responses will sometimes do so because S is close to zero, but on other occasions will do so because a positive value of S (that is, an internal representation that has non-zero strength) is combined with a negative value of the error term. By selecting trials *post hoc* in which the participant says 'invisible', the experimenter relinquishes control over the items included in the analysis, in which this bias may therefore intrude. If it were possible to obtain a second independent awareness report for these items, regression to the mean in the error term would pull these reports upwards and turn many of them into 'visible' reports, as in the continuous case illustrated in Figure 1C. The onus must be on the researcher to eliminate this possibility, for example, by showing that her visibility scale has very high reliability.

As an illustration of how this issue undermines research findings, consider a recent neuroimaging study by Sheikh et al. (2019) whose research question was whether the brain extracts semantic information from unconscious (masked) words. On each trial a word was briefly flashed and participants made two responses: first, they judged whether it was animate or inanimate, and secondly they reported their conscious experience of the word on a 3-point scale from “I didn't see anything” (i.e., invisible) to “I think I saw the word clearly or almost clearly”. Sheikh et al.’s main finding was that activity in several brain regions distinguished animate from inanimate words even when they were judged invisible. This inference rests, however, on the unlikely assumption that the visibility reports are perfectly reliable and free of random error (it also pays no heed to the flexibility with which the participant sets her response criterion; Schmidt, 2015). Any tendency for error to contaminate these reports must inevitably mean that amongst the trials classified as invisible are some (perhaps many) where the true latent state driving these reports is greater than zero: they are, in effect, misclassified[3]. Shanks (2017) provides simulations of this type of situation where the awareness report is discrete. There is simply no getting away from the ineluctable effects of regression to the mean.

Sheikh et al. (2019) provided a further reason to believe that awareness was truly absent on ‘invisible’ trials: objective awareness, as measured by the accuracy of animate/inanimate judgments, was close to and not significantly greater than chance (50%) on these trials. But resonating with our earlier discussion, the upper limit of the 95% confidence interval (CI) of this estimate was ~70%. Hence the data look much more like the pattern shown in Figure 1B than in Figure 1A. This study therefore illustrates not only the unwarranted use of trial-based *post hoc* selection but also the dangers of slipping from failure to reject the null to acceptance of the null.

Loken and Gelman (2017) pointed out that studies (particularly ones with small samples) affected by measurement error may erroneously yield statistically significant effects. The magnitude of these effects will inevitably shrink in replications in which these errors regress. Applied to find evidence of unconscious processes, the *post hoc* selection method represents the flip side of this coin, namely where measurement error leads to erroneous null results (the absence of awareness). It is a method that cannot fulfil the hope of researchers who use it in the expectation that it can reveal true unconscious processes.

[3] Of course it will also be the case that some words judged as visible will in fact be truly invisible, but this has no bearing on the interpretation of brain activity on trials judged invisible.

Correlation and regression slope

If the analytic strategies discussed above fall short of providing the evidence needed to infer unconscious processing, then perhaps stronger inferences can be drawn from the pattern of correlation observed between performance and awareness? Looking at Model 1 it should be clear from the equations in Table 1 that they can be rearranged to yield the simple linear equation

$$P = c \cdot A - c \cdot e_A + e_P, \qquad (1)$$

where $c = a/b$. Hence the model predicts a non-zero correlation between performance and awareness so long as $a > 0$ and $b \neq 0$. The only circumstances in which this is not the case are when performance is at chance (that is, when $a = 0$) or awareness reports are completely random ($b = 0$). This seems intuitively reasonable. If a common latent variable determines both performance and awareness, then they must be measurably correlated so long as performance (and hence S) is greater than chance and awareness reports have some degree of validity. The degree of this correlation may vary depending on how much error there is in the two measures, but nonetheless the correlation cannot be zero.

It is clear that many researchers implicitly accept the above reasoning, because correlations have been regularly (and perhaps increasingly) used to justify claims about unconscious processes. Malejka et al. (2021, Table 1) provide examples of the use of correlation analyses to infer unconscious processes in response inhibition, contextual cuing, evaluative and classical conditioning, artificial grammar learning, language and category learning, and other domains. For example, Salvador et al. (2018) recently reported a pair of experiments on memory suppression. The experimental details are not important here, except to note that they measured awareness in a discrimination test for briefly flashed shapes previously presented as cues in a memory suppression task. Although discrimination was significantly better than chance (that is, conforming to the pattern shown in Figure 1B), awareness was uncorrelated ($r = -.09$, combined across the two experiments) with performance (a memory suppression dependent measure).

Like the other approaches we have discussed, this method raises a number of issues. To begin with, just because a correlation is low and not significantly different from zero does not entail that the true relationship is zero – a conclusion that cannot easily be drawn if one is employing Null Hypothesis Significance Testing (NHST). In Salvador et al.'s (2018) data, the very low power of their experiments meant that the 95% CI on their correlation estimate was [-.31, .15]. This wide interval is compatible with a nontrivial positive correlation. This problem could be addressed by using substantially larger samples, allowing the CI to be more narrowly estimated, but another problem would remain. It goes without saying that the measures Salvador et al. obtained, like all measures in behavioural research, were subject to measurement error. In the presence of such error, a correlation coefficient will invariably underestimate the true relationship. This regression attenuation or regression dilution phenomenon (Spearman, 1910) has been much studied in areas such as biomedical statistics (see Hutcheon et al., 2010). In economics the underestimation of relationships in regression models when the predictors contain error variance is so firmly established that it has been termed an 'iron law' (Hausman, 2001).

To estimate the degree of this attenuation in Salvador et al.'s (2018) data, Malejka et al. (2021) computed the reliability of their performance and awareness measures via odd–even split-half correlations. Remarkably, the resulting value for the awareness test was a mere .15. It was much higher (.65) for the performance measure. [As an aside, while both reliabilities matter in correlation analyses, the reliability of the predictor variable – awareness in this case – is much more important when performance is being regressed onto awareness, for reasons explained by Hutcheon et al. (2010)]. The reliability of Salvador et al.'s awareness measure is far below the minimum level expected of a psychometrically-sound measure. Collapsing data across both experiments, the Pearson correlation coefficient between performance and awareness when disattenuated for measurement error was -.28 with an extremely wide 95% CI of [-1.00, .47]. Clearly, based on the measures they used, Salvador et al. are only able to rule out correlations larger than about .5. Stated differently, their finding of a nonsignificant performance-awareness correlation only challenges single-process models that predict fairly high correlations.

One of the key points of the current article (see Box 1) is that future research in this field must report the reliabilities of the measures obtained (particularly measures of awareness). Readers might be sceptical of this need, perhaps thinking that the Salvador et al. (2018) reliabilities are atypical. In fact there is an accumulating body of evidence that they are far from unusual, and indeed we recommend a default assumption that standard measures used in this research field have medium-to-low reliabilities (LeBel & Paunonen, 2011). Studies that have reported reliabilities have usually revealed disappointingly low values (Geyer et al., 2020; Kalra et al., 2019; Vadillo et al., 2020; West et al., 2018).

Comprehensive data illustrating this point come from very large samples analyzed by Vadillo et al. (in press). These data are from contextual cuing experiments using a popular visual search task (see Goujon et al., 2015). On each of many visual search trials, participants locate a target (often an inverted T) amongst distractors (Ls) as rapidly as possible. Some of the distractor patterns repeat during the experiment, always with the target in the same location, and response times (RTs) confirm that participants learn about these repeating displays as their responses become progressively faster, relative to novel patterns. After many blocks of search trials, participants' awareness of the repeating patterns is probed. One common assessment of awareness comprises a recognition test in which they are explicitly instructed to judge whether displays are repeated or novel. In another type of test, displays are shown in which the target has been replaced by another distractor, and participants are instructed to point to the location of the hidden target. Vadillo et al. collected data from 505 participants and also analyzed data from three experiments by Colagiuri and Livesey (2016) with sample sizes of 63, 84 and 766. The methods, which are standard across dozens of contextual cuing studies, were quite similar across these experiments.

Vadillo et al. (in press) found that the split-half reliability of the performance measure in these studies was very consistently slightly less than .5 (with Spearman-Brown correction). The reliability of the awareness measures was around .5 for the experiments conducted by Vadillo et al. but closer to .3 for Colagiuri and Livesey's experiments. While these figures are not quite as starkly low as those in Salvador et al.'s experiments, noted above, there are good reasons to suspect that they overestimate the reliabilities of these measures in similar studies, which usually include much briefer (and therefore noisier) tests of awareness. In any case, these figures confirm the considerable magnitude of measurement error in these scales as well as the impossibility of obtaining high performance-awareness correlations: if a measure does not correlate well with itself – which is what reliability measures – it is unlikely to correlate any better with another measure[4].

With this very large sample, a meta-analysis yielded a mean performance-awareness correlation of .039, [-.014, .092]. One might look at this as providing quite strong support for independence between the measures and as suggesting that the true correlation must be close to zero. But a further meta-analysis which took account of the reliabilities yielded a substantially larger disattenuated mean correlation of r = .104 as well as a much wider CI [-.016, .225] – an example of the 'iron law' in action. Hence even with data from over 1400 participants and a nonsignificant correlation, the hypothesis that the correlation is as large as .225 cannot be rejected. To put this in perspective, it is larger than the average effect size obtained across all psychological research! In an analysis of 200 meta-analyses, Stanley et al. (2018) estimated the mean effect size at r = .19.

In the face of these problems, what can researchers do? A powerful approach is to switch from NHST and instead conduct a Bayesian analysis that permits us to quantify the degree of support for the null compared to the alternative (r > 0) hypothesis. Malejka et al. (2021) described 3 such Bayesian models. The reader is referred to that article for details. A Bayesian approach requires a specification of the prior distribution of correlations which could for example be a uniform distribution (all correlations between 0 and 1 are equally plausible *a priori*), or, more plausibly, one in which small positive correlations are more likely. For present purposes, Malejka et al.'s key finding in their reanalysis of Salvador et al.'s data was that the Bayes Factor (BF) in favour of the null hypothesis was only in the region of 1.8 when a prior was used in which small positive correlations are assumed to be more likely. Thus on a plausible assumption that the true correlation is positive but small, Salvador et al.'s data only marginally favoured the null over the alternative hypothesis.

More constructively, Malejka et al. used these Bayesian models as tools for deriving sample size predictions which researchers can use to guide their sample size planning. Given pilot work to enable the reliability of the experimental measures to be estimated, together with a specification of the minimum true correlation one wishes to be able to rule out, Malejka et

---

[4] To be precise, the upper bound on the X-Y correlation is the product of the square roots of the reliabilities of X and Y.

al.'s simulations specify the sample sizes needed to achieve a given level of evidence (e.g., BF > 10 in favour of the null). For instance, if the true correlation is 0 and reliability is above .65 for both measures, then a commonly achievable sample size of 300 is adequate to find median BFs > 10 in favour of the null (in comparison, there were 44 and 30 participants in Salvador et al.'s experiments).

Regression intercept

The above discussion highlights the serious interpretive consequences of measurement error in correlational analyses. Such error leads to systematic underestimation of the true correlation, and we have seen evidence to suggest that in some popular tasks for studying unconscious mental processes, this error is sufficiently large (reliability is low) for the effect on measured correlation to be non-trivial. A second and quite common type of analytic strategy that researchers sometimes apply to performance/awareness correlational data proves to be equally problematic, albeit for a different reason. This is the practice of focusing attention on the regression intercept. Why is the intercept an interesting measure?

Since the two error terms in Equation 1 both have means of 0, it seems natural to infer that the predicted value of P when $A = 0$ is also zero, which seems intuitive given that this model only includes a single latent variable. Hence it is very tempting to assume that any single-process model must predict an intercept of zero: when performance is regressed onto awareness, the former must be zero when the latter is.

This logic is widely employed. For instance, in a recent study by Berkovitch and Dehaene (2019) participants judged whether a series of target words were nouns or verbs, each of which was preceded by a masked subliminal prime (also a noun or a verb). The results showed a typical compatibility effect, with faster responses to target words preceded by primes of the same grammatical category. Prime awareness was based on participants' ability to discriminate the primes as nouns or verbs in a visibility task conducted at the end of the experiment. Berkovitch and Dehaene noted that "the intercept of this regression was significant [...], indicating that priming remained significant even at null *d*'" (p. 31). Greenwald et al. (1995) were the first to propose this analytic method, claiming that "the finding of a significant regression intercept effect (i.e., a nonzero indirect effect associated with a zero direct effect) indicates an unconscious contribution to the indirect measure" (p. 26).

But once again this is a fallacious conclusion (Miller, 2000; Sand & Nilsson, 2016). Regression dilution not only flattens the slope of a regression line, it also systematically biases the intercept upwards. This can be seen in Figure 1D. Model 1B, which differs from Model 1A simply in the parameter values, was used to generate the data in this figure. Thus once again the performance and awareness measures are linked to a common latent construct (*S*). From Equation 1 we can calculate that the true slope at the latent level is $c = a/b = 100/0.3 = 333$, shown by the black line in Figure 1D, whereas the empirical regression line (in pink in Figure 1D) has a slope far shallower than this (an increase in awareness of 1 unit is associated with an increase in performance of about 30 units). This is the dilution effect discussed above. But what we can also see is that the observed intercept is not zero, it is much higher than this. Hence measurement error leads to systematic bias in the estimation of the regression intercept. Error-free measurement of the predictor variable, an assumption of standard regression, is unlikely to be achievable in any behavioural research. Interested readers can readily run this simulation from the openly-available R code to see that it is error in the awareness measure (e_A) that is crucial to this effect. When this error value is set to 0, the true intercept (0) is estimated. This is not the case when error (e_P) in the outcome variable (performance) is set to zero (Hutcheon et al., 2010).

Why does this effect occur? For exactly the same reason as for the *post hoc* selection case, namely regression to the mean. From the model we can see that because *S* has a mean and standard deviation of 1.0, for most participants its value is greater than zero. This implies that when a given participant's measured awareness is at or below 0, it will often be because, by chance, e_A is negative. When we calculate this participant's performance, it will be based on the same value of *S* but combined with e_P which is statistically independent of e_A. The former has an expected value of 0, resulting in a performance score that regresses towards the group mean. Measured performance is closer to the group mean performance level than measured awareness is to group mean awareness.

To amplify, if we take all of the simulated participants in Figure 1D whose awareness is close to zero (let's say between -0.1 and 0.1 on the awareness scale), we can ask what the true level of awareness is in these participants. The answer is that it is very close to 0.3, the mean awareness of the group as a whole (because b = 0.3 in this simulation). This is very similar to the situation in Figure 1C, discussed previously. For these participants, the awareness error term e_A is on average about -0.3, thus cancelling out the positive true awareness[5]. The positive regression intercept is an inevitable consequence: because *S* is positive rather than zero in these participants, then when combined with independent error to determine their performance level, the latter is robustly positive. Figure 1D depicts a pattern in which average awareness seems barely above chance, performance is clearly above chance overall, the correlation between performance and awareness is very low, and the level of performance when awareness is at chance is also above chance. Yet all of these features emerge from a model in which a single latent process determines both performance and awareness.

We would see exactly the same but in reverse if instead of regressing *P* onto *A* we regressed *A* onto *P*. We would now find that participants with a mean performance level of zero have above-chance awareness! Those scoring at zero on the performance measure will usually only do so because e_P is large and negative. By the same argument as above but turned around, their level of awareness will be close to the group mean awareness because e_A, being independent of e_P, will be close to zero. In both cases the effect is not indicative of some separable latent process, it is a statistical artifact caused by regression to the mean.

To be fair, Greenwald and colleagues were sufficiently aware of this concern that they sought to address it directly. They proposed a 'bias correction' method intended to provide an accurate estimate of the true regression intercept, taking the effects of measurement error into account (Klauer et al., 1998). In one recent application, Greenwald and De Houwer (2017) estimated the relationship between awareness and conditioning. In several experiments they observed intercepts greater than zero (as well as conditioning-awareness correlations not significantly greater than zero), and reported that they remained positive even when corrected.

[5] As before, this is an extreme example designed to illustrate a point. If the relative amount of error in the measurement of awareness were smaller, the true level of awareness in these participants would not be so close to the group mean (0.3), though it would still be greater than 0.

Although it is rarely used, the development of this correction method is an important contribution. However it is unclear whether it can solve the underlying problem, because it is relatively easy to construct situations in which it under-corrects (Miller, 2000). We (Malejka et al., 2019) generated data from a single-system model similar to Model 1 and then corrected the intercept estimates using the Klauer et al. (1998) method. Despite the fact that the true intercept was known to be zero, the method falsely concluded that it was positive in a high proportion of cases, depending on the precise parameters of the simulation. For instance, the under-correction was particularly acute when the underlying data included only a small proportion of truly unaware participants. Until a more robust correction method is available (see Hutcheon et al., 2010, for references to the statistical literature on this problem), we urge researchers to be extremely cautious in using of the regression intercept method to draw inferences about unconscious processes (Box 1, Recommendation 3).

Single dissociations

Many reported performance/awareness dissociation patterns are essentially claims about correlation, and hence the major observations in the previous section apply to them. For example, several studies have included explicit instructions to orient participants towards the key stimulus feature or regularity, confirmed that these instructions boost awareness, and then shown that the performance measure is unaffected by these instructions (e.g., Jiménez et al., 1996; Westerberg et al., 2011). In essence this is the claim that there is no correlation (across groups) between performance and awareness under the instructional manipulation. The famous claim that hippocampal brain damage affects explicit memory while leaving implicit memory unaffected (e.g., Hamann & Squire, 1997) is again a claim about a null correlation.

Because single-process models, as shown above, can predict correlations that are so low they are hard to distinguish from zero, these patterns of single dissociation are generally not diagnostic of distinct latent processes for performance and awareness (for examples, see Berry et al., 2008; Jamieson et al., 2010; Nosofsky & Zaki, 1998; Shanks & Berry, 2012). It would only be in the case that the measures are highly reliable (Box 1, Recommendation 3) that such patterns would pose a serious challenge to models incorporating a single latent process.

Conclusions

Inferring unconscious mental processes from patterns of bivariate performance/awareness data is fraught with under-appreciated problems. In this article we have based our discussion of these problems around some extremely simple models (Table 1), but it is our strong belief that reflecting on the predictions of such models can illuminate the valid inferences that can be drawn from bivariate data patterns. Just to give one example, our informal probing (via problems like that in Box 2) reveals that many researchers believe that random error is equally distributed across the range of a measured variable. This is false – at the positive extreme of the measured variable, errors will on average be positive and at the negative extreme they will be negative, and this is why regression dilution by measurement error occurs. The implications of this simple statistical fact for the interpretation of bivariate performance/awareness data are profound, as explained in our discussion of *post hoc* selection and the interpretation of correlational data. Other strong but false intuitions (for instance, that a model with a single latent process necessarily predicts a zero intercept) can equally lead to incorrect conclusions about unconscious mental processes.

These considerations lead to the Recommendations listed in Box 1. Within the domain of bivariate data, we advise less use of some common practices (e.g., relying on regression intercepts) and greater use of others (e.g., calculating confidence intervals or Bayes Factors). What else can researchers do? Studies that include experimental manipulations of awareness and performance are important. Any demonstration, for example, that an independent variable increases performance at the same time as reducing awareness (that is, a double dissociation) would be extremely hard to explain with reference to only a single latent construct. Further efforts to induce strong dissociations of this sort are likely to be illuminating. For example, there have been demonstrations of manipulations that increase priming whilst decreasing prime visibility (e.g., Biafora & Schmidt, 2020), and of other manipulations that affect conditioned responses and conscious expectancies in opposite direction (Perruchet, 2015).

However we also urge researchers to extend beyond this domain of methods and analytic approaches. Ultimately, bivariate patterns are not particularly constraining in relation to possible latent factors. In the Introduction we pointed out that decades of developments in psychometrics have yielded a considerably broader set of methods and tools for inferring mental processes. Research on the Implicit Association Test (Kurdi et al., 2019; Nosek & Smyth, 2007) provides on exemplar of this very different strategy. Wider adoption of these approaches – which would require employing multiple awareness and performance tests and comparison of single- versus multiple-process models – might considerably strengthen the inferences we can draw about unconscious processes.

References

Abadie, M., & Waroquier, L. (2019). Evaluating the benefits of conscious and unconscious thought in complex decision making. *Policy Insights from the Behavioral and Brain Sciences*, *6*, 72-78. https://doi.org/10.1177/2372732218816998

Beauny, A., de Heering, A., Muñoz Moldes, S., Martin, J.-R., de Beir, A., & Cleeremans, A. (2020). Unconscious categorization of sub-millisecond complex images. *PLoS One*, *15*, Article e0236467. https://doi.org/10.1371/journal.pone.0236467

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*, 1293-1295.

Berkovitch, L., & Dehaene, S. (2019). Subliminal syntactic priming. *Cognitive Psychology*, *109*, 26-46. https://doi.org/10.1016/j.cogpsych.2018.12.001

Berry, C. J., Kessels, R. P. C., Wester, A. J., & Shanks, D. R. (2014). A single-system model predicts recognition memory and repetition priming in amnesia. *Journal of Neuroscience*, *34*, 10963-10974. https://doi.org/10.1523/jneurosci.0764-14.2014

Berry, C. J., Shanks, D. R., & Henson, R. N. A. (2008). A single-system account of the relationship between priming, recognition, and fluency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 97-111. https://doi.org/10.1037/0278-7393.34.1.97

Berry, C. J., Ward, E. V., & Shanks, D. R. (2017). Does study duration have opposite effects on recognition and repetition priming? *Journal of Memory and Language*, *97*, 154-174. https://doi.org/10.1016/j.jml.2017.07.004

Biafora, M., & Schmidt, T. (2020). Induced dissociations: Opposite time courses of priming and masking induced by custom-made mask-contrast functions. *Attention, Perception, & Psychophysics*, *82*, 1333-1354. https://doi.org/10.3758/s13414-019-01822-4

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*, 663-677. https://doi.org/10.1177/0956797620915887

Blake, R. (2021). Reflections on Eriksen's seminal essay on discrimination, performance and learning without awareness. *Attention, Perception, & Psychophysics*, *83*, 546-557. https://doi.org/10.3758/s13414-020-02098-9

Campbell, D. T., & Kenny, D. A. (1999). *A primer on regression artifacts.* Guildford.

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

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*, 1-61.

Cohen, M. A., Ostrand, C., Frontero, N., & Pham, P.-N. (2021). Characterizing a snapshot of perceptual experience. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000864

Colagiuri, B., & Livesey, E. J. (2016). Contextual cuing as a form of nonconscious learning: Theoretical and empirical analysis in large and very large samples. *Psychonomic Bulletin & Review*, *23*, 1996–2009. https://doi.org/10.3758/s13423-016-1063-0

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302. https://doi.org/10.1037/h0040957

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F., & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395*, 597-600.

Dienes, Z. (2015). How Bayesian statistics are needed to determine whether mental states are unconscious. In M. Overgaard (Ed.), *Behavioural methods in consciousness research* (pp. 199-220). Oxford University Press.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, *69*, 338 - 351.

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS One*, *7*, e29081.

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review*, *95*, 91-101.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, *67*, 279-300.

Geyer, T., Rostami, P., Sogerer, L., Schlagbauer, B., & Müller, H. J. (2020). Task-based memory systems in contextual-cueing of visual search and explicit recognition. *Scientific Reports*, *10*, Article 16527. https://doi.org/10.1038/s41598-020-71632-4

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, *19*, 524-533. https://doi.org/10.1016/j.tics.2015.07.009

Grahek, I., Schaller, M., & Tackett, J. L. (2021). Anatomy of a psychological theory: Integrating construct-validation and computational-modeling methods to advance theorizing. *Perspectives on Psychological Science*, *16*, 803-815. https://doi.org/10.1177/1745691620966794

Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconceiving the relation between conscious and unconscious. *American Psychologist*, *72*, 861-871. https://doi.org/10.1037/amp0000238

Greenwald, A. G., & De Houwer, J. (2017). Unconscious conditioning: Demonstration of existence and difference from conscious conditioning. *Journal of Experimental Psychology: General*, *146*, 1705-1721. https://doi.org/10.1037/xge0000371

Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: Dissociations of unconscious from conscious cognition. *Journal of Experimental Psychology: General*, *124*, 22-42. https://doi.org/10.1037/0096-3445.124.1.22

Hamann, S. B., & Squire, L. R. (1997). Intact perceptual memory in the absence of conscious memory. *Behavioral Neuroscience*, *111*, 850-854.

Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science*, *8*, 195-207. https://doi.org/10.1177/1745691612460684

Hausman, J. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives*, *15*, 57-67. https://doi.org/10.1257/jep.15.4.57

Hedger, N., Gray, K. L. H., 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*, 934-968. https://doi.org/10.1037/bul0000054

Heycke, T., Aust, F., & Stahl, C. (2017). Subliminal influence on preferences? A test of evaluative conditioning for brief visual conditioned stimuli using auditory unconditioned stimuli. *Royal Society Open Science*, *4*, Article 160935. https://doi.org/doi:10.1098/rsos.160935

Hutcheon, J. A., Chiolero, A., & Hanley, J. A. (2010). Random measurement error and regression dilution bias. *BMJ*, *340*, c2289. https://doi.org/10.1136/bmj.c2289

Jamieson, R. K., Holmes, S., & Mewhort, D. J. K. (2010). Global similarity predicts dissociation of classification and recognition: Evidence questioning the implicit-explicit learning distinction in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1529–1535.

Jiang, Y. V., Sha, L. Z., & Sisk, C. A. (2018). Experience-guided attention: Uniform and implicit. *Attention, Perception, & Psychophysics*, *80*, 1647-1653. https://doi.org/10.3758/s13414-018-1585-9

Jiménez, L., Méndez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 948-969.

Kalra, P. B., Gabrieli, J. D. E., & Finn, A. S. (2019). Evidence of stable individual differences in implicit learning. *Cognition*, *190*, 199-211. https://doi.org/10.1016/j.cognition.2019.05.007

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*, 321-340. https://doi.org/10.1016/j.cognition.2010.05.011

Kellen, D., Davis-Stober, C. P., Dunn, J. C., & Kalish, M. L. (2021). The problem of coordination and the pursuit of structural constraints in psychology. *Perspectives on Psychological Science*, *16*, 767-778. https://doi.org/10.1177/1745691620974771

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review*, *110*, 728-744. https://doi.org/10.1037/0033-295x.110.4.728

Klauer, K. C., Draine, S. C., & Greenwald, A. G. (1998). An unbiased errors-in-variables approach to detecting unconscious cognition. *British Journal of Mathematical and Statistical Psychology*, *51*, 253-267.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahnik, S., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., Cemalcilar, Z., Chandler, J., Cheong, W., Davis, W. E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., . . . Nosek, B. A. (2014). Investigating variation in replicability: A "many labs" replication project. *Social Psychology*, *45*, 142-152. https://doi.org/10.1027/1864-9335/a000178

Konstantinidis, E., & Shanks, D. R. (2014). Don't Bet on It! Wagering as a Measure of Awareness in Decision Making Under Uncertainty. *Journal of Experimental Psychology-General*, *143*(6), 2111-2134. https://doi.org/10.1037/a0037977

Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., & Banaji, M. R. (2019). Relationship between the implicit association test and intergroup behavior: A meta-analysis. *American Psychologist*, *74*, 569-586. https://doi.org/10.1037/amp0000364

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2018). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B*, *75*, 45-57. https://doi.org/10.1093/geronb/gby065

Le Pelley, M. E., Newell, B. R., & Nosofsky, R. M. (2019). Deferred feedback does not dissociate implicit and explicit category learning systems: Commentary on Smith et al. (2014). *Psychological Science*, *30*, 1403-1409. https://doi.org/10.1177/0956797619841264

LeBel, E. P., & Paunonen, S. V. (2011). Sexy but often unreliable: The impact of unreliability on the replicability of experimental findings with implicit measures. *Personality and Social Psychology Bulletin*, *37*, 570-583.

LeDoux, J. E., Michel, M., & Lau, H. (2020). A little history goes a long way toward understanding why we study consciousness the way we do today. *Proceedings of the National Academy of Sciences*, *117*, 6976-6984.

https://doi.org/10.1073/pnas.1921623117

Lodder, P., Ong, H. H., Grasman, R. P. P. P., & Wicherts, J. M. (2019). A comprehensive meta-analysis of money priming. *Journal of Experimental Psychology: General*, *148*, 688-712.

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, *355*, 584-585. https://doi.org/10.1126/science.aal3618

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology-Animal Behavior Processes*, *28*(1), 3-26. https://doi.org/10.1037//0097-7403.28.1.3

Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences*, *102*, 16075-16080.

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*, Article 104667.

Malejka, S., Vadillo, M. A., & Shanks, D. R. (2019). *Pitfalls in the use of regression analysis to investigate unconscious conditioning.* International Convention of Psychological Science, Paris.

Merikle, P. M., & Reingold, E. M. (1992). Measuring unconscious perceptual processes. In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness: Cognitive, clinical, and social perspectives.* Guilford Press.

Mertens, G., & Engelhard, I. M. (2020). A systematic review and meta-analysis of the evidence for unaware fear conditioning. *Neuroscience & Biobehavioral Reviews*, *108*, 254-268. https://doi.org/10.1016/j.neubiorev.2019.11.012

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

Miller, J. (2000). Measurement error in subliminal perception experiments: Simulation analyses of two regression methods. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1461-1477.

Morton, V., & Torgerson, D. J. (2005). Regression to the mean: Treatment effect without the intervention. *Journal of Evaluation in Clinical Practice*, *11*, 59-65. https://doi.org/10.1111/j.1365-2753.2004.00505.x

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, *5*, 309-315.

Newell, B. R., & Dunn, J. C. (2008). Dimensions in data: Testing psychological models using state-trace analysis. *Trends in Cognitive Sciences*, *12*, 285-290.

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, *37*, 1-19. https://doi.org/10.1017/s0140525x12003214

Neyman, J., & Pearson, E. S. (1933). IX. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London A*, *231*, 289–337. https://doi.org/10.1098/rsta.1933.0009

Nobre, A. P., Melo, G. M., Gauer, G., & Wagemans, J. (2020). Implicit processing during inattentional blindness: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *119*, 355-375. https://doi.org/10.1016/j.neubiorev.2020.10.005

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test. *Experimental Psychology*, *54*, 14-29. https://doi.org/10.1027/1618-3169.54.1.14

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, *9*, 247-255.

O'Donnell, M., Nelson, L. D., Ackermann, E., Aczel, B., Akhtar, A., Aldrovandi, S., Alshaif, N., Andringa, R., Aveyard, M., Babincak, P., Balatekin, N., Baldwin, S. A., Banik, G., Baskin, E., Bell, R., Bialobrzeska, O., Birt, A. R., Boot, W. R., Braithwaite, S. R., . . . Zrubka, M. (2018). Registered replication report: Dijksterhuis and van Knippenberg (1998). *Perspectives on Psychological Science*, *13*, 268-294. https://doi.org/10.1177/1745691618755704

Peirce, C. S., & Jastrow, J. (1884). On small differences of sensation. *Memoirs of the National Academy of Sciences*, *3*, 75-83.

Perruchet, P. (2015). Dissociating conscious expectancies from automatic link formation in associative learning: A review on the so-called Perruchet effect. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*, 105-127. https://doi.org/10.1037/xan0000060

Phillips, I. (2021). Blindsight is qualitatively degraded conscious vision. *Psychological Review*, *128*, 558–584. https://doi.org/10.1037/rev0000254

Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*, 291-321.

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

Röseler, L., Schütz, A., Blank, P. A., Dück, M., Fels, S., Kupfer, J., Scheelje, L., & Seida, C. (2021). Evidence against subliminal anchoring: Two close, highly powered, preregistered, and failed replication attempts. *Journal of Experimental Social Psychology*, *92*, Article 104066. https://doi.org/10.1016/j.jesp.2020.104066

Salvador, A., Berkovitch, L., Vinckier, F., Cohen, L., Naccache, L., Dehaene, S., & Gaillard, R. (2018). Unconscious memory suppression. *Cognition*, *180*, 191-199. https://doi.org/10.1016/j.cognition.2018.06.023

Sand, A., & Nilsson, M. E. (2016). Subliminal or not? Comparing null-hypothesis and Bayesian methods for testing subliminal priming. *Consciousness and Cognition*, *44*, 29-40. https://doi.org/10.1016/j.concog.2016.06.012

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

Schimmack, U. (2021). The Implicit Association Test: A method in search of a construct. *Perspectives on Psychological Science*, *16*, 396-414. https://doi.org/10.1177/1745691619863798

Schmidt, T. (2015). Invisible stimuli, implicit thresholds: Why invisibility judgments cannot be interpreted in isolation. *Advances in Cognitive Psychology*, *11*, 31-41. https://doi.org/10.5709/acp-0169-3

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge: Cambridge University Press.

Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review*, *24*(3), 752-775. https://doi.org/10.3758/s13423-016-1170-y

Shanks, D. R., & Berry, C. J. (2012). Are there multiple memory systems? Tests of models of implicit and explicit memory. *Quarterly Journal of Experimental Psychology*, *65*(8), 1449-1474. https://doi.org/10.1080/17470218.2012.691887

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367-395. https://doi.org/10.1017/S0140525X00035032

Shanks, D. R., Vadillo, M. A., Riedel, B., Clymo, A., Govind, S., Hickin, N., Tamman, A. J. F., & Puhlmann, L. M. C. (2015). Romance, risk, and replication: Can consumer choices and risk-taking be primed by mating motives? *Journal of Experimental Psychology: General*, *144*, E142-E158. https://doi.org/10.1037/xge0000116

Sheikh, U. A., Carreiras, M., & Soto, D. (2019). Decoding the meaning of unconsciously processed words using fMRI-based MVPA. *NeuroImage*, *191*, 430-440. https://doi.org/10.1016/j.neuroimage.2019.02.010

Sisk, C. A., Remington, R. W., & Jiang, Y. V. (2020). A spatial bias toward highly rewarded locations is associated with awareness. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *46*, 669-683. https://doi.org/10.1037/xlm0000749

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, *3*, 271–295. https://doi.org/10.1111/j.2044-8295.1910.tb00206.x

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin*, *144*, 1325-1346. https://doi.org/10.1037/bul0000169

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*, Article e3001241. https://doi.org/10.1371/journal.pbio.3001241

Stein, T., Utz, V., & van Opstal, F. (2020). Unconscious semantic priming from pictures under backward masking and continuous flash suppression. *Consciousness and Cognition*, *78*, Article 102864. https://doi.org/10.1016/j.concog.2019.102864

Stephens, R. G., Matzke, D., & Hayes, B. K. (2019). Disappearing dissociations in experimental psychology: Using state-trace analysis to test for multiple processes. *Journal of Mathematical Psychology*, *90*, 3-22. https://doi.org/10.1016/j.jmp.2018.11.003

Vadillo, M. A., Hardwicke, T. E., & Shanks, D. R. (2016). Selection bias, vote counting, and money-priming effects: A comment on Rohrer, Pashler, and Harris (2015) and Vohs (2015). *Journal of Experimental Psychology: General*, *145*, 655-663. https://doi.org/10.1037/xge0000157

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, *23*(1), 87-102. https://doi.org/10.3758/s13423-015-0892-6

Vadillo, M. A., Linssen, D., Orgaz, C., Parsons, S., & Shanks, D. R. (2020). Unconscious or underpowered? Probabilistic cuing of visual attention. *Journal of Experimental Psychology: General*, *149*, 160-181. https://doi.org/10.1037/xge0000632

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

Ward, E. V., Berry, C. J., Shanks, D. R., Moller, P. L., & Czsiser, E. (2020). Aging predicts decline in explicit and implicit memory: A life-span study. *Psychological Science*, *31*, 1071-1083. https://doi.org/10.1177/0956797620927648

Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature*, *217*, 972-974.

Weiskrantz, L. (1997). *Consciousness lost and found: A neuropsychological exploration.* Oxford University Press.

West, G., Shanks, D. R., & Hulme, C. (2021). Sustained Attention, Not Procedural Learning, is a Predictor of Reading, Language and Arithmetic Skills in Children. *Scientific Studies of Reading*, *25*(1), 47-63. https://doi.org/10.1080/10888438.2020.1750618

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Developmental Science*, *21*, Article e12552. https://doi.org/10.1111/desc.12552

Westerberg, C. E., Miller, B. B., Reber, P. J., Cohen, N. J., & Paller, K. A. (2011). Neural correlates of contextual cueing are modulated by explicit learning. *Neuropsychologia*, *49*, 3439-3447.

Wildegger, T., Myers, N. E., Humphreys, G., & Nobre, A. C. (2015). Supraliminal but not subliminal distracters bias working memory recall. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 826-839.

Yeates, F., Wills, A. J., Jones, F. W., & McLaren, I. P. L. (2015). State trace analysis: Dissociable processes in a connectionist network? *Cognitive Science*, *39*, 1047-1061.

Zaki, S., Nosofsky, R. M., Jessup, N. M., & Unverzagt, F. W. (2003). Categorization and recognition performance of a memory-impaired group: Evidence for single-system models. *Journal of the International Neuropsychological Society*, *9*, 394-406.

Table 1

Models used to generate the data in Figure 1.

<table>
<tr><th>Model</th><th>Performance</th><th>Awareness</th><th>Parameters</th></tr>
<tr><td>Model 2</td><td>$P = a \cdot S_P + e_P$</td><td>$A = b \cdot S_A + e_A$</td><td>$S_P \sim N(1.0, 0.2)$
$a = 100$
$e_P \sim N(0, 80)$
$S_A \sim N(0, 0.1)$
$b = 1.0$
$e_A \sim N(0, 0.1)$</td></tr>
<tr><td>Model 1a</td><td rowspan="2">$P = a \cdot S + e_P$</td><td rowspan="2">$A = b \cdot S + e_A$</td><td>$S \sim N(1.0, 0.2)$
$a = 100$
$e_P \sim N(0, 80)$
$b = 1.0$
$e_A \sim N(0, 1.0)$</td></tr>
<tr><td>Model 1b</td><td>$S \sim N(1.0, 1.0)$
$a = 100$
$e_P \sim N(0, 80)$
$b = 0.3$
$e_A \sim N(0, 1.0)$</td></tr>
</table>

Note: Model names indicate whether 1 or 2 latent processes are included. The normal distribution *N*() is parameterized by its mean and standard deviation (*SD*).

Box 1

Recommendations for future research using bivariate data to infer unconscious mental processes.

	Recommendation
1	We recommend the wider use of frequentist-statistical approaches with fixed error rates and Bayesian statistical approaches, to avoid slippage from failure to reject the null hypothesis (awareness = 0) to asserting that the null is true. Bayes Factors permit evaluation of the strength of evidence in support of the null.
2	Sample sizes should be planned a priori to allow for either high-powered tests (in the frequentist approach) or clear support for one of the competing hypotheses (in the Bayesian approach) on (a) the average measure of awareness and (b) the performance-awareness correlation.
3	Researchers should report the reliabilities of their dependent measures, and especially of their awareness measure, and attempt to increase these reliabilities when they are inadequate. When these measures are based on multiple trials, a split-half correlation can be calculated without any additional data needing to be collected. When awareness is based on a single response, a retest should be included in the experimental design.
4	We advise extreme caution regarding the use of the *post hoc* selection and regression intercept methods, which rest on assumptions that are often unwarranted (e.g., negligible influence of measurement error).
5	Eliciting trial-by-trial awareness reports (e.g., 'visible'/'invisible') and analyzing a performance measure just across the invisible trials – a form of *post hoc* selection – should only be undertaken when the awareness test can be demonstrated to have high reliability.

Box 2

Imagine that you are a food chemist asked by a supermarket to measure the toxicity of 20 food additives, select the 5 least toxic ones, and report the average toxicity of these 5 additives.

You measure the toxicity of all 20 additives, but your equipment cannot measure toxicity perfectly. Your measurements are inevitably noisy and include some random measurement error. This error is independent of true toxicity.

Having selected the 5 least toxic ones, you average their measured toxicities and report this figure to the supermarket.

Is there any problem with this procedure?

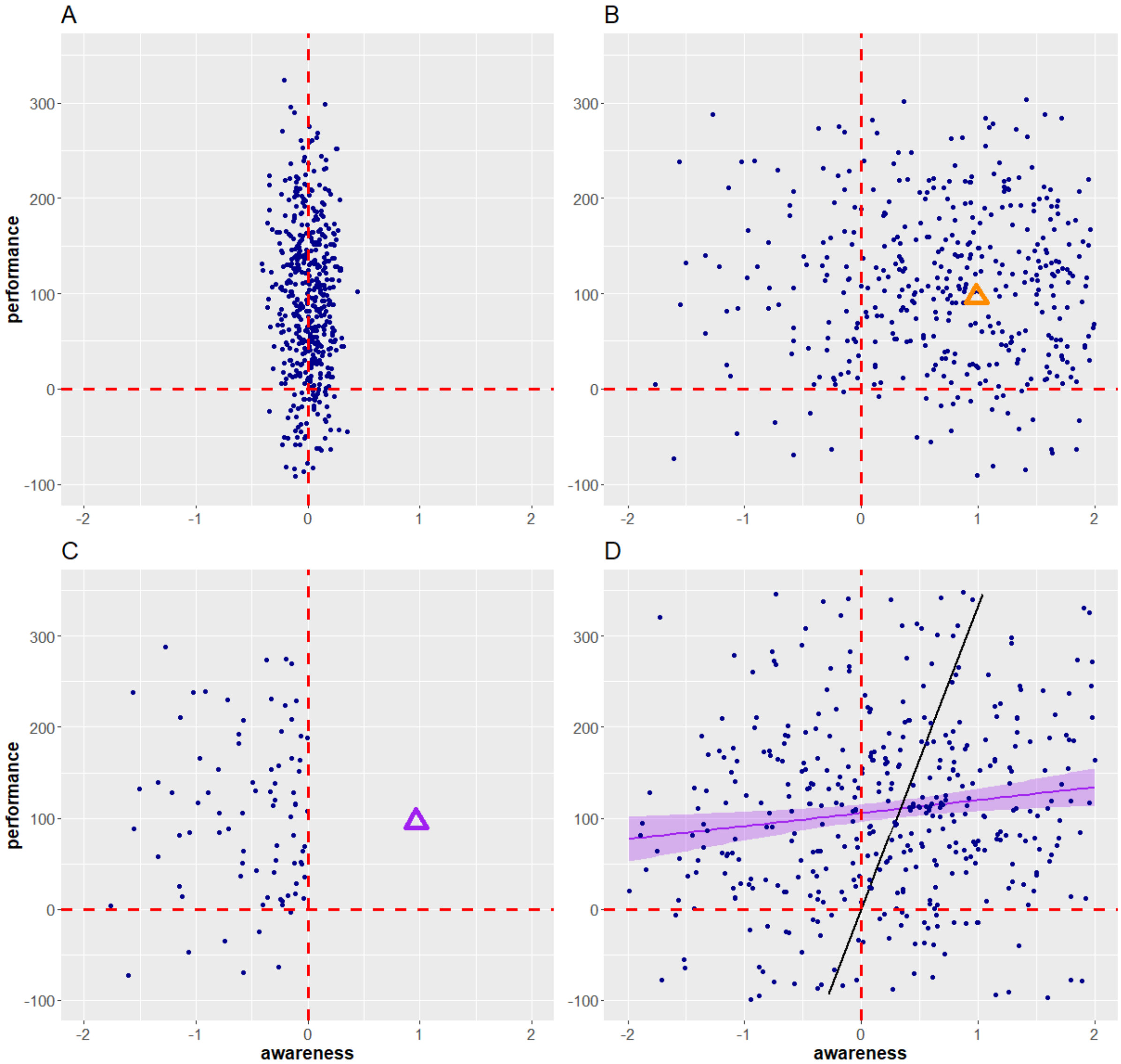

Figure 1

Example data patterns. A: Data from Model 2. Awareness scores have a true mean close to zero. B: Data from Model 1a. Awareness scores have a true mean greater than zero. The orange triangle is the mean. C: Same as B but data from all participants with awareness scores greater than 0 have been removed. The purple triangle shows the observed mean performance and true mean awareness of the included participants. D: Data from Model 1b. The best-fitting regression line (pink) shows a weak correlation between performance and awareness. The shaded area shows the 95% confidence interval around the regression line. The black line shows the true relationship at the latent level.