

Self-Reports Are Better Measurement Instruments than Implicit Measures

Olivier Corneille¹ & Bertram Gawronski²

¹ Université Catholique de Louvain, Louvain-la-Neuve, Belgium

² University of Texas at Austin, Austin, Texas, USA

Self-report measures directly ask respondents to report their mental content such as thoughts and feelings. By contrast, implicit measures aim to assess thoughts and feelings via performance indicators (for example, response times, error rates and response frequencies) under conditions that favor automatic processing. Implicit measures are now widely used within psychological science and beyond because they are assumed to be superior to self-reports in various regards. In this Perspective, we argue that, despite the enthusiasm for implicit measures, self-reports are most often the better measurement option. First, the use of implicit measures is often based on mistaken assumptions about unique disadvantages of self-reports. Second, self-reports have advantageous characteristics that are currently unmatched in implicit measures. We call for a more sophisticated use of self-reports and for caution when using implicit measures in basic and applied research.

Introduction

Social scientists have traditionally relied on self-reports for measuring people's thoughts and feelings. In self-report measurement, respondents are asked to introspect and report on information directly relevant to the construct of interest. For example, in research on life satisfaction participants might be asked to rate how happy they are with their life; in research on consumer preferences participants might be asked to rate how much they like different soda brands; and in research on gender stereotyping participants might be asked to rate men and women in terms of gender stereotypical traits.

Although self-report measures are commonly used, they have been criticized for several reasons, two of which have received particular attention. First, respondents might be unwilling to provide accurate self-reports.^{1,2} For example, participants with low life satisfaction might report a higher level of satisfaction to deceive themselves or the researcher about their actual satisfaction. Second, respondents might be unable to provide accurate self-reports because they lack conscious access to the relevant thoughts and feelings.^{3–5} For example, participants might have no idea how they feel about their life, leading them to respond in a manner that does not reflect their actual life satisfaction.

Indirect behavioral measures in which participants do not have to introspect and report on their thoughts and feelings have been used to overcome these issues.^{6,7} Over the past three decades, interest in a new generation of such measures, often referred to as implicit measures, has flourished in psychology⁸ as well as other disciplines (for example, political science^{9–11}, marketing^{12–14}, education^{15–17}, and economics^{18–20}). Today, there are more than two dozen implicit measures, including the implicit association test,²¹ the affect misattribution procedure,²² and the evaluative priming task.²³ These measures have been

used to assess a variety of constructs such as attitudes,²¹ stereotypes,²⁴ self-esteem,²⁵ motives,^{26,27} outcome expectations,²⁸ values,²⁹ self-harm ideation,^{30,31} and semantic relations between concepts.^{32,33} The implicit association test specifically has been applied to countless domains, including perceptions of the U.S. Supreme Court³⁴ and combat veterans³⁵, attitudes about carbon storage³⁶, associations between disability and assistance dogs³⁷, alcohol-related associations³⁸ and risk of self-harm.³⁰

Although the tasks commonly classified as implicit measures differ in terms of their procedural details and the cognitive processes involved,^{39,40} they share two features. First, mental contents are inferred from performance indicators (for example, response times, error rates or response frequencies) typically measured via computers. Second, responses on implicit measures are assumed to be driven by automatic processes (processes that are unintentional, unconscious, efficient, or uncontrolled). Based on these characteristics, implicit measures are often thought to provide unique information (for example, information about unconscious mental contents) or superior information (for example, information uncontaminated by self-presentation motives) relative to self-reports. In some cases, these assumptions can lead researchers to miss or disregard information they collected on self-reports despite the potential value of this information for the research objectives. For example, the authors of a study that failed to replicate effects of exposure to counterstereotypical exemplars on an implicit measure of racial prejudice did not relate the replication failure to the finding that self-reports of racial prejudice had robustly reversed in their study samples⁴¹ relative to the original sample.⁴²

In this Perspective, we challenge the enthusiasm about implicit measures and argue that self-reports almost always remain the better measurement option. First, we review claims about unique disadvantages of self-reports and the evidence that they are either

unfounded or also apply to implicit measures. Next, we describe unmatched advantages that self-reports have over implicit measures. We then argue that self-reports are better than implicit measures for almost all research questions addressed by implicit measures and consider potential objections to our arguments. We conclude by calling for a more sophisticated use of self-reports and greater caution about the frequently proclaimed—yet empirically questionable—advantages of implicit measures.

Purported Disadvantages of Self-Reports

In this section, we discuss six widespread—yet unfounded—claims about unique disadvantages of self-reports.

Contamination

Self-reports are thought to be contaminated by external influences, including context, social desirability, and experimental demands. By contrast, implicit measures are often assumed to be “bona-fide pipelines” that “get inside the head” of respondents in a nonreactive fashion,²³ therefore providing uncontaminated information about mental contents.

There are justified concerns that contextual factors can influence responses on self-reports.^{43–45} For example, correlations between self-reports of general happiness and self-reported frequency of dating are higher when the dating question precedes (vs. follows) the happiness question.⁴⁶ Importantly, responses on implicit measures are also malleable and highly sensitive to contextual influences (for reviews see ref ^{47–49}). However, these findings have had little impact on the modal narrative in the literature on implicit measures.⁵⁰

Concerns about social-desirability effects (a tendency to answer questions in a manner that will be viewed favourably by others) on self-reports are also very common. However, the available evidence suggests that these concerns might be overblown. In fact, full anonymity has been found to decrease rather than increase the accuracy of self-reports by decreasing accountability and increasing survey satisficing (taking ‘short cuts’ to complete a survey faster).⁵¹ Critically, implicit measures are not immune to social-desirability effects. For example, responses to minority groups on implicit measures are more favorable when the research assistant is a member of the focal minority group.^{52,53}

A weaker claim is that implicit measures are less sensitive to social desirability compared to self-report measures. However, the empirical basis for this claim is not clear. For example, one study found that public (versus private) testing conditions reduced scores indicative of anti-gay prejudice on a gay/straight implicit association test to a larger extent ($d = .40$) than self-reported homophobia ($d = .17$) and heterosexism

($d = .35$).⁵⁴ Although the mechanisms underlying these effects are still unclear, the findings indicate that responses on implicit measures depend on the social desirability of their expression in the testing situation.

Moreover, if implicit measures are immune or less sensitive to social desirability effects they should converge less with self-reports in socially sensitive domains. However, evidence for this assertion is inconclusive, as findings differ across studies,^{55–58} across analytic approaches within studies,⁵⁷ and across categories of tasks within studies.⁵⁶ Likewise, one might expect smaller correlations between self-reports and implicit measures among individuals who show greater social desirability concerns, but this has not been robustly observed⁵⁶. Thus, at present there is no compelling evidence that implicit measures are immune or less sensitive than self-reports to social desirability effects.

A related concern about self-reports is that they are susceptible to experimental demand effects. The idea underlying this concern is that participants might draw inferences about the experimental hypotheses and adjust their self-reports to confirm these hypotheses. A common assumption is that demand effects might contaminate self-reports but not implicit measures because responses on the latter are less controllable.^{59–61} Although demand effects have the potential to contaminate self-reports via conscious and less conscious control strategies,⁶² it is unclear why implicit measures should be immune to such influences.

In fact, implicit measures are notoriously sensitive to faking strategies.^{53,63–67} Social biases on implicit measures are lower when participants are told that their “job is to be as nonprejudiced as possible” compared to when they are not.⁵³ Implicit measures are also more sensitive to counter-attitudinal information when participants are invited under hypnotic suggestion to let this information “sink in more deeply than is typically the case” compared to when they are not.⁶⁸ Thus, contrary to uncontrollability claims, responses on implicit measures can be controlled in response to direct requests or mere suggestion communicated by the experimenter in the testing situation. Alignment between direct or inferred experimental requests and responses on implicit measures might be implemented through control strategies other than faking, such as mental imagery (for example, imagining a woman with counterstereotypical characteristics⁶⁹ or imagining a false alibi⁷⁰), or phenomenological control, in which individuals implement a behavior in a controlled way while remaining unaware that they are doing so.^{62,71}

Consciousness

Another common claim about implicit measures is that they provide access to unconscious thoughts and feelings that cannot be captured with self-report measures. This idea is most prominently reflected in

the claim that implicit measures uncover hidden biases that people do not know they have.^{72–74} It is true that direct self-report measures require awareness of the to-be-measured thoughts and feelings. However, the claim that implicit measures capture unconscious thoughts and feelings is a theoretical hypothesis,⁷⁵ and the available empirical evidence suggests that it is incorrect.^{76,77} For example, counter to the idea that implicit measures capture unconscious thoughts and feelings, participants are able to predict their scores on implicit measures with a high level of accuracy (correlations between predicted and actual scores greater than .50)^{78–80} regardless of how much prior experience they have with implicit measures, how much information they receive about implicit measures, and whether implicit measures are described to participants as measures of ‘true beliefs’ or ‘cultural associations’⁷⁸. Prediction accuracy is not perfect, allowing for the possibility that the thoughts and feelings captured by implicit measures are at least partially unconscious. However, there is no empirical evidence supporting the latter interpretation; in the absence of such evidence it seems more likely that less-than perfect prediction accuracy is the product of noisy measurement.

Accuracy of implicit measurement score predictions is higher at a within-subjects level (within-subject correlations between predicted and actual scores greater than .50) involving comparisons across different target objects (for example, a participant’s liking of apples compared to oranges) than at a between-subjects level (between-subjects correlations between predicted and actual scores in the range of .20 to .30) involving comparisons across different participants (for example, a participant’s liking of apples compared to the liking of apples by other participants).^{78,81} Accurate predictions at a between-subjects level require not only knowledge of one’s own standing on the relevant dimension (for example, how much one likes apples), but also knowledge about the relative standing of the other participants (for example, how much the other participants in the sample like apples).⁷⁷ Thus, higher prediction accuracy for within-subjects versus between-subjects comparisons is consistent with the fact that participants have accurate knowledge of their own preferences but limited knowledge about the preferences of other participants. Lack of knowledge about the preferences of other people has arguably little to do with unconsciousness and therefore low prediction accuracy at a between-subjects level cannot be interpreted as evidence for lack of awareness.⁸²

It is often assumed that participants express surprise in response to feedback about their social biases on the implicit association test because they lack awareness of these biases.^{83,84} 9/30/2024 7:20:00

PMHowever, these commonly observed surprise reactions can be explained as the product of statistical artifacts that inflate measurement scores and arbitrary conventions in the description of these scores.^{77,85,86}

Thus, although it is true that self-report measures are suboptimal for capturing unconscious thoughts and feelings, there is no evidence that implicit measures capture unconscious thoughts and feelings either. Importantly, prominent authors who initially promoted the hypothesis that implicit measures provide access to unconscious thoughts and feelings have acknowledged its inaccuracy.^{87,88} Nevertheless, the original hypothesis continues to spread and shape educational and intervention programs.^{89,90}

Non-automaticity

It is often assumed that self-reports are ill-suited to study automatic processes, and that implicit measures are superior for that purpose. Although it is true that implicit measures involve constraints that favour automatic processes, responses on self-reports are also shaped by automatic processes, especially under conditions conducive to automatic processing. For example, researchers interested in uncontrolled mental processes could collect self-reports while instructing participants to avoid specific influences when providing their self-report response (for example, to avoid being influenced by the attractiveness of job candidates when judging their suitability for a position);⁹¹ researchers interested in fast mental processes could request speeded self-reports;⁹² and researchers interested in unconscious mental processes could request that participants judge a stimulus only if they feel they cannot report context information that was originally associated with it.⁹³

Process dissociation is a particularly valuable approach to disentangling automatic and non-automatic influences on self-reports (Box 1).^{94,95} A central feature of process-dissociation procedures is the comparison of responses across conditions in which automatic and non-automatic processes influence responses in the same direction (compatible condition) and conditions in which automatic and non-automatic processes influence responses in the opposite direction (incompatible condition). By comparing self-reports across compatible and incompatible conditions, the respective contributions of automatic and non-automatic processes can be quantified mathematically.^{94,95} Process-dissociation procedures have been successfully used to study uncontrolled^{96–98} and unconscious (that is, memory-independent)^{99–103} influences on self-reports.

Failure to capture simple associations

It is commonly assumed that responses on self-reports reflect propositional beliefs (mental propositions about relations between objects and events that are deemed true or false), whereas

responses on implicit measures reflect simple associations (mental associations between objects and events independent of their relation and perceived truth; see also below).¹⁰⁴ Although the existence of simple mental associations has been disputed,¹⁰⁵ simple associations (unlike propositional beliefs) are insensitive to truth and relational meaning. For example, being exposed to the proposition ‘it is not true that vaccines cause autism’ might create a simple association between the concepts ‘vaccines’ and ‘autism’. Likewise, being exposed to the proposition ‘aspirin reduces headaches’ might create a simple association between the concepts ‘aspirin’ and ‘headaches’. These simple associations are assumed to elicit negative reactions to vaccines and aspirin, respectively, even when people understand and accept the positive implications of the relevant proposition. It has been claimed that responses on implicit measures reflect simple associations irrespective of truth and relational meaning, and that such associations cannot be captured with self-report measures.

There are two problems with these assumptions. First, counter to the claim that responses on implicit measures reflect simple associations, meta-analytic evidence suggests that implicit measures are highly sensitive to truth and relational meaning.¹⁰⁶ Second, research using process-dissociation procedures suggests that responses on self-report measures are jointly shaped by propositional beliefs and simple associations.^{107–109} In fact, effects of simple associations on self-reports have been found even under conditions where responses on implicit measures reflected only propositional beliefs without any detectable effect of simple associations.^{107,110} Thus, the available evidence suggests that self-reports (combined with process-dissociation procedures) are superior in capturing simple associations than implicit measures.

Robustness

According to a prominent theoretical view, there is a direct link between associative learning and responses on implicit measures³⁹, such that mental associations build up slowly as a result of repeated exposure to co-occurrences between stimuli.^{111–113} When these associative links are sufficiently strong, they are assumed to be difficult to change, making responses on implicit measures highly robust and resistant to novel information.¹¹⁴ However, these views are challenged by studies showing that responses on implicit measures can change quickly in response to minimal information, including a single statement.^{115–117} Moreover, although some changes on implicit measures are transient,¹¹⁸ there is no evidence that changes on implicit and self-report measures differ in terms of their relative stability over time at the individual level. A similar conclusion applies to studies

that examined changes over large time windows at the societal level, which have generally found similar trajectories for explicit and implicit measures of attitudes.^{119–121} Thus, again, the presumed disadvantage of self-reports relative to implicit measures is unsubstantiated.

Person vs. environment

Responses on implicit measures are traditionally interpreted as reflecting characteristics of the person. Alternative theorizing suggests that responses on implicit measures are more informative about aspects of the person’s environment.¹²² Such interpretations are prominent in research on prejudice and stereotyping, where implicit measures have been claimed to capture systemic rather than personal biases.¹²³ In research inspired by this view, data from participants who reside in the same region are aggregated, and analyses test whether these aggregate scores are systematically related to characteristics of the region. For example, studies have shown that regional differences in racial bias measured by the implicit association test are associated with regional differences in race-related disparities in police traffic stops¹²⁴, the use of lethal force in policing¹²⁵, and mortality rates.¹²⁶ Findings like these gave rise to the idea that implicit measures have unique value for studying systemic biases.¹²⁷

However, with few exceptions, aggregated responses on self-report and implicit measures at the regional level tend to be redundant, with correlations as high as $r = .94$ (ref¹²⁸). In these cases, regional aggregates of self-report and implicit measures almost always show identical results,¹²⁸ and some studies even found stronger effects for self-report than implicit measures.^{126,129,130} Thus, self-report measures typically produce the same results as implicit measures, and implicit measures therefore do not provide unique insights into systemic biases.

Advantages of Self-Reports

In the previous section, we discussed six widespread—yet unfounded—claims about unique disadvantages of self-reports. In this section, we discuss three reasons why self-reports should be considered superior to implicit measures.

Reliability

Self-report measures tend to be more reliable on average than implicit measures. Of the currently available implicit measures, the implicit association test and the affect misattribution procedure are the only two instruments that show estimates of internal consistency that meet standard psychometric criteria with Cronbach’s alphas between .70 and .90 (ref^{40,88}). All other instruments suffer from substantial measurement error with modal Cronbach’s alphas around .50. Indeed, Cronbach’s alphas are often lower

than that and sometimes close to zero.^{40,88} Although internal consistency must be evaluated on a case-by-case basis, established self-report measures typically fare much better in this regard. For example, commonly used self-report measures of social biases typically show Cronbach's alphas around .80 (refs^{131–133}). Moreover, even the two implicit measures that have shown acceptable estimates of internal consistency (the implicit association test and the affect misattribution procedure) show lower test-retest stability ($r_s = \sim .50$) than conceptually corresponding self-report measures ($r_s = \sim .75$).¹³⁴

The differential reliability of self-report and implicit measures has substantial implications for basic and applied research. First, low internal consistency of a given measure reduces statistical power for identifying existing effects with that measure.^{135,136} Second, low temporal stability of a given measure reduces statistical power for identifying cross-temporal relations with other measures. Overall, the higher internal consistency and test-retest stability of self-report measures compared to implicit measures makes them superior for use in basic and applied research.

Prediction

Self-report measures are also generally better for predicting behavior than implicit measures. Meta-analyses on the predictive validity of implicit and self-report measures have shown that both perform similarly, with behavioral criteria showing weak to modest correlations with implicit measures ($r = .13$ to $r = .28$) and self-report measures ($r = .11$ to $r = .36$)^{57,137–139}. However, these meta-analyses focused exclusively on studies that included both self-reports and implicit measures, thereby ignoring a large set of studies that included only self-report measures. This limitation is important, because studies with implicit measures rarely rely on established theoretical frameworks (such as the theory of planned behavior¹⁴⁰) or the measurement approaches suggested by these frameworks for predicting behavior, which might impact predictive validity. Indeed, meta-analyses on studies guided by these frameworks generally show stronger associations between self-report measures and behavioral criteria ($r = .40$ to $r = .50$) than meta-analyses comparing the predictive validity of implicit measures and self-reports ($r = .11$ to $r = .36$)^{141–144}, with meta-analytic correlations as high as $r = .58$ between newly formed self-reported behavioral attitudes and actual behavior.^{145,146} These results suggest that, when all available meta-analytic evidence is taken into account, self-report measures are superior for predicting behavior compared to implicit measures.

Theoretically oriented researchers have pointed out that although self-report measures might outperform implicit measures in predicting deliberate behavior, implicit measures might outperform self-

report measures in predicting spontaneous behavior.^{104,147} At the measurement level, correlations between a predictor variable and a behavioral criterion should be greater when the processing conditions during the measurement of the predictor variable correspond to the processing conditions during the measurement of the behavioral criterion (for example, both involve fast responses).⁸ Thus, given that implicit measures involve constraints that favor effects of automatic processes and self-report measures involve constraints that favor effects of non-automatic processes, implicit measures might be expected to outperform self-report measures in predicting behavior guided by automatic processes (spontaneous behavior) whereas self-report measures might be expected to outperform implicit measures in predicting behavior guided by non-automatic processes (deliberate behavior)^{104,148}.

Although these assumptions might be correct for traditional applications of self-report measures that do not involve any processing constraints^{57,147} the presumed superiority of implicit measures is based on the misconception that self-report measures are ill-suited for capturing responses arising from automatic processes. However, it is possible to use self-report measures to capture responses arising from automatic processes⁹². Thus, to the extent that the presumed superiority of implicit measures for predicting spontaneous behavior is rooted in their sensitivity to automatic processes and self-report measures can be adapted to quantify the impact of both automatic and non-automatic processes, self-report measures might be deemed superior to implicit measures because they can be adapted to predict both spontaneous and deliberate behavior. This conclusion also applies to cases where implicit measures contribute to predicting behavior over and above traditional self-report measures.^{137,138,149} In such cases, self-report measures that capture effects of automatic processes should contribute to predicting behavior over and above self-report measures that capture effects of non-automatic processes, analogous to the incremental contribution of implicit measures. However, the latter conclusion is admittedly speculative owing to the lack of relevant data.

Finally, a well-established advantage of self-report measures for predicting behavior is their higher temporal stability compared to implicit measures.¹³⁴ This issue is highly relevant for research that aims to predict behavior over time: it is unrealistic to expect substantial correlations between a predictor measure and a behavioral criterion over time if the predictor measure shows low correlations with itself over time.¹⁵⁰ Although the low test-retest stability of implicit measures can be overcome via repeated measurement to reduce error noise (for example,

asking participants to complete the same implicit association test five times in a single session),¹⁵¹ such a strategy seems impractical and difficult to implement in many studies. Aggregating data from repeated measurements also increases the overlap between self-report and implicit measures,¹⁵¹ raising the question of why one should administer the same implicit measure multiple times when the same information can be obtained with a simple self-report measure—especially when self-report measures based on a single item can reasonably predict ($r = \sim .15$) important life events (for example, job change or relationship dissolution) one year ahead of time.¹⁵² These considerations corroborate the conclusion that self-report measures are generally superior to implicit measures for predicting behavior.

Flexibility

A final advantage of self-reports over implicit measures is their flexibility. Self-reports can be used to assess thoughts and feelings about complex issues that are difficult to gauge with implicit measures (for example, thoughts about a proposed policy change). Furthermore, when collected in the context of open questions or semi-structured interviews, self-reports enable respondents to voice their lived experiences in a richer and more flexible way than implicit measures.¹⁵³ Implicit measures typically impose impoverished and constrained communication formats, reducing the expression of complex subjective experiences to binary button presses. On a structural level, self-reports do not have to assume, impose, and reproduce dichotomous contrasts between social categories (for example, Asian vs. white; male vs. female; old vs. young), and indeed such dichotomies seem highly problematic in light of the growing body of work on intersectionality.¹⁵⁴

Finally, the lower level of technology required by self-report measures compared to implicit measures makes them more inclusive for both researchers and respondents. For instance, interviews or questionnaire measures can be more easily implemented in the field than computer-based measures, and they are also better suited to respondents with low digital literacy or impaired motor functions.

Implications

Based on the above considerations, we conclude that self-report measures are most often a better choice than implicit measures for addressing the kinds of research questions that are commonly studied with implicit measures. When combined with process-dissociation procedures, self-report measures are superior to implicit measures even in areas where implicit measures have been claimed to provide unique insights that cannot be gained from self-report measures, such as effects of simple mental associations.

Although implicit measures can still be valuable for addressing automaticity questions, claims about automaticity cannot be drawn from analyzing findings with implicit measures alone, because the same findings might be obtained with a measure that involves non-automatic processing. Moreover, when drawing inferences about automaticity features (awareness, intention, efficiency, or control) from differences between measures, it is critical that the to-be-compared measures differ only in terms of the focal automaticity feature, and that this difference is not confounded with other factors. This criterion renders direct comparisons between self-reports and implicit measures problematic because the two kinds of measures differ in numerous ways (for example, stimuli and response formats) beyond the presumed difference in automaticity.^{150,155} Structural-fit paradigms are better for studying automaticity features than comparisons between self-reports and implicit measures that differ in multiple ways beyond the focal automaticity feature.

In structural-fit paradigms, responses are compared across two variants of the same task that differ only with regard to the automaticity feature of interest.^{92,155–157} For example, to study the role of intentionality (vs. unintentionality), researchers could use two variants of the affect misattribution procedure, in which participants are presented with attitudinal prime stimuli (for example, a photo of a person), followed by neutral Chinese ideographs (Figure 1a).¹⁵⁵ In the regular version of the task, participants might be asked to evaluate the Chinese ideographs and ignore prime stimuli.²² If the evaluations of the Chinese ideographs systematically differ as a function of the preceding face primes (for example, if the ideographs are evaluated more favorably following younger than older face primes), then this procedure provides an indirect assessment of participants' responses to the primes. In the modified version, participants might be asked to directly evaluate the prime stimuli and ignore the Chinese ideographs, providing a direct assessment of participants' responses to the prime stimuli.¹⁵⁵ Differences in evaluative responses (measured indirectly or directly) to the prime stimuli across the two task variants can be interpreted as reflecting differences between unintentional (in the regular task variant) and intentional (in the modified task variant) evaluations of the prime stimuli.¹⁵⁵

Although the affect misattribution procedure provides an ideal paradigm to study intentionality with a structural-fit approach, it seems difficult (if not impossible) to apply a structural-fit approach to other implicit measures. These difficulties limit the value of these instruments for studying features of automaticity, which is ironically their most frequently proclaimed advantage over self-report measures. Indeed, for

automaticity features other than intentionality (which can be studied with two variants of the affect misattribution procedure, Figure 1a), applying a structural-fit approach to self-report measures might be the better choice (Figure 1b). For example, to study resource-dependence, researchers could compare responses on self-report measures under conditions where participants have to perform a difficult secondary task (for example, rehearsing an eight-digit letter string) versus an easy secondary task (for example, rehearsing a two-digit letter string); to study controllability, researchers could compare responses on self-report measures under conditions where participants are instructed versus not instructed to avoid some unwanted influence on their judgments; to study awareness, researchers could compare responses on self-report measures under conditions of fast versus slow stimulus presentation.

As with any method, validation work is necessary to confirm the reliability of process-dissociation and structural-fit approaches when addressing automaticity questions. In process-dissociation procedures, this can be achieved by testing whether parameter estimates are influenced by experimental manipulations^{97,99} or predict behaviors^{97,99} in a theoretically sound way. More validation work of this sort is still needed.³⁹ It should also be noted that methodological advances in self-report research go beyond process-dissociation and structural-fit approaches. For example, in attitude research self-reports of felt ambivalence have been used to address questions about controllability¹⁵⁸ and simple associations.¹⁰⁹

Potential Objections

A potential objection to our conclusions about the limited value of implicit measures is that self-report and implicit measures are weakly correlated with each other,⁵⁵ and have distinct antecedents¹⁵⁹ and behavioral correlates.¹⁴⁷ One could argue that these dissociations between self-reports and implicit measures provide a strong justification for including implicit measures in any study with traditional self-reports because they suggest that implicit measures provide unique insights that cannot be gained from self-report measures.

We agree that dissociations between self-report and implicit measures can provide an interesting starting point for future research, but we urge caution in interpreting such dissociations in terms of inherent differences between the two kinds of measures (for example, as implied by claims that self-report and implicit measures capture related but distinct constructs). Self-report and implicit measures differ in numerous aspects, which renders dissociations between the two difficult to interpret. For example, in research on racial bias, it remains unclear whether dissociations between measures are driven by the

nature of the measurement instruments (self-report vs. implicit) or the different stimulus materials (for example, presenting faces in the implicit but not self-report measure).¹⁵⁰ Although research using a structural-fit approach has replicated some previously obtained dissociations between self-report and implicit measures,¹⁵⁵ other dissociations disappeared when confounds between measurement type and procedural factors were controlled.¹⁵⁶ Thus, when studying potential dissociations, researchers should heed the lesson of the structural-fit approach by comparing responses across two variants of the same task that differ only with regard to the critical feature of interest.

Another potential objection is that implicit measures might provide valuable information when performance indicators (for example, error rates or response times) are subjected to computational modeling, including applications of process dissociation and multinomial processing tree modeling.¹⁶⁰ To the extent that implicit measures can be understood as experimentally controlled paradigms that mirror the processes underlying spontaneous responses more broadly¹⁶¹, combining them with computational modeling could provide nuanced insights into spontaneous behavior that cannot be gained from either self-reports or traditional applications of implicit measures.

However, an important caveat is that broader conclusions from research using computational models to analyze responses on implicit measures are based on the premise of process correspondence (the assumption that the processes underlying spontaneous responses on implicit measures correspond to the processes underlying spontaneous behavior in natural settings).^{8,40} Applications of computational modeling to implicit measures would have little value for understanding spontaneous behavior in natural settings if spontaneous behavior in natural settings involves different processes from those underlying responses on implicit measures. Research ignoring the premise of process correspondence could be criticized for its narrow focus on understanding responses on a specific task and for losing sight of the more important question of what drives spontaneous behavior in natural settings.

A final objection is that some of our arguments against implicit measures refer to lack of evidence, which implies that our conclusions might have to be revised in light of new findings. We agree that the lack of supportive evidence is not the same as counterevidence.¹⁶² However, it is concerning for a research program that has inspired such a large number of studies over a period of almost three decades to defend major assumptions with lack of counterevidence rather than supportive evidence. Although novel applications of process-dissociation

and structural-fit paradigms require thorough validation just like any other method, we think it is time to acknowledge the limits of implicit measures and that self-report measures are a better and more promising measurement option.

Conclusions

Over the past three decades, psychologists have devoted considerable effort to the development of implicit measures. This research program has flourished based on assumptions about disadvantages of self-reports relative to implicit measures. We argued that six core assumptions justifying this measurement revolution remain unsubstantiated or have been proven wrong. Our discussion of contamination suggests that responses on both self-report and implicit measures are sensitive to context, social-desirability, and experimental-demand effects. Overturning these conclusions would require dismissing a large body of evidence showing that implicit measures are sensitive to these effects. Our conclusions regarding unconsciousness are similarly difficult to reject considering evidence indicating that people can predict the outcomes of implicit measures.¹⁶³ Echoing our conclusion on this point, prominent implicit social cognition researchers have toned down their earlier claims that implicit measures capture mental contents that respondents are not aware of^{87,88}, and a consensus seems to be emerging that intentionality is a more relevant automaticity feature instead.^{164,165}

We also argued that questions of automaticity cannot be reliably addressed using only implicit measures, or by comparing them to tasks that differ in various ways beyond the focal automaticity feature. Finding robust dissociations on structurally matched self-report and implicit measures would support the value of the latter. However, structural-fit paradigms for implicit measures have only been developed for the affect misattribution procedure, and it is currently unclear how structurally matched versions of other implicit measures could be designed.

Moreover, there is strong evidence that self-reports can capture simple associations, and it is not clear why this evidence should be dismissed. The same argument applies to our dismissal of the robustness hypothesis, which has been debunked in numerous studies. Finally, overturning our conclusions regarding the person vs. environment debate would require robust evidence indicating that, at least for some topics, implicit measures provide additional insights relative to self-reports. Needless to say this demonstration should use structurally matched measures so that the findings can be interpreted unambiguously.

We also discussed three advantages of self-reports over implicit measures: their greater reliability, predictive value, and flexibility. Challenging our

conclusions regarding reliability would require new implicit measures or task variants that show strong psychometric qualities, including high construct validity.^{166,167} Dozens of implicit measures have thus far failed to achieve this goal, and these efforts might have distracted researchers from using more advanced self-report measures. Regarding prediction, overturning our conclusion would require new implicit measures that have higher predictive value than self-reports, again ideally achieved using structurally-matched measures. Finally, it is currently unclear how implicit measures could become more flexible and inclusive than self-reports.

Although the impetus for developing and using implicit measures three decades ago was understandable, and this research program brought about valuable insights, we believe the time is ripe for taking stock and acknowledging that, in the end, self-reports most often do a better measurement job, both from a methodological and a pragmatic standpoint. We therefore call for enhanced and more sophisticated reliance on self-reports in psychological science and in related scientific domains, including for automaticity questions that could leverage structural-fit and process-dissociation procedures.

Acknowledgements

The authors thank J  r  my B  na, Mayan Navon, and Jessica R  hner for helpful comments on a previous version of this article.

Author Contributions

The authors contributed equally to all aspects of the article.

Competing Interests

The authors declare no competing interests.

References

1. Lalwani, A. K., Shavitt, S. & Johnson, T. What is the relation between cultural orientation and socially desirable responding? *J. Pers. Soc. Psychol.* **90**, 165–178 (2006).
2. Paulhus, D. L. Two-component models of socially desirable responding. *J. Pers. Soc. Psychol.* **46**, 598–609 (1984).
3. Wilson, T. D. *Strangers to ourselves: Discovering the Adaptive Unconscious*. (Belknap, Cambridge, Mass London, 2004).
4. Greenwald, A. G. & Banaji, M. R. Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychol. Rev.* **102**, 4–27 (1995).
5. Nisbett, R. E. & Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. *Psychol. Rev.* **84**, 231–259 (1977).
6. Webb, E. J., Campbell, D.T., Schwartz, R.D.,

- & Sechrest, L. *Unobtrusive measures*. (Sage, Thousand Oaks, 2000).
7. Campbell, D. T. The indirect assessment of social attitudes. *Psychol. Bull.* **47**, 15–38 (1950).
 8. Gawronski, B., De Houwer, J. & Sherman, J. W. Twenty-five years of research using implicit measures. *Soc. Cogn.* **38**, s1–s25 (2020).
 9. Gawronski, B., Galdi, S. & Arcuri, L. What can political psychology learn from implicit measures? Empirical Evidence and New Directions. *Polit. Psychol.* **36**, 1–17 (2015).
 10. Ryan, T. J. & Krupnikov, Y. Split feelings: Understanding implicit and explicit political persuasion. *Am. Polit. Sci. Rev.* **115**, 1424–1441 (2021).
 11. Clayton, K., Horrillo, J. & Sniderman, P. M. The BIAT and the AMP as measures of racial prejudice in political science: A methodological assessment. *Polit. Sci. Res. Methods* **11**, 363–373 (2023).
 12. Dimofte, C. V. Implicit measures of consumer cognition: A review. *Psychol. Mark.* **27**, 921–937 (2010).
 13. Slabbinck, H. & Spruyt, A. On the Selection and Use of Implicit Measures in Marketing Research: A Utilitarian Taxonomy. in *Review of Marketing Research* (eds. Baumgartner, H. & Weijters, B.) 171–210 (Emerald Publishing Limited, 2022). doi:10.1108/S1548-643520220000019010.
 14. Brunel, F. F., Tietje, B. C. & Greenwald, A. G. Is the Implicit Association Test a Valid and Valuable Measure of Implicit Consumer Social Cognition? *J. Consum. Psychol.* **14**, 385–404 (2004).
 15. Glock, S. & Kovacs, C. Educational Psychology: Using Insights from Implicit Attitude Measures. *Educ. Psychol. Rev.* **25**, 503–522 (2013).
 16. Denessen, E., Hornstra, L., Van Den Bergh, L. & Bijlstra, G. Implicit measures of teachers' attitudes and stereotypes, and their effects on teacher practice and student outcomes: A review. *Learn. Instr.* **78**, 101437 (2022).
 17. Pit-ten Cate, I. M. & Glock, S. Teachers' Implicit Attitudes Toward Students From Different Social Groups: A Meta-Analysis. *Front. Psychol.* **10**, 2832 (2019).
 18. Stanley, D. A., Sokol-Hessner, P., Banaji, M. R. & Phelps, E. A. Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proc. Natl. Acad. Sci.* **108**, 7710–7715 (2011).
 19. Alesina, A., Carlana, M., La Ferrara, E. & Pinotti, P. Revealing stereotypes: Evidence from immigrants in schools. *Am. Econ. Rev.* **114**, 1916–1948 (2024).
 20. Carlana, M. Implicit Stereotypes: Evidence from Teachers' Gender Bias. *Q. J. Econ.* **134**, 1163–1224 (2019).
 21. Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. Measuring individual differences in implicit cognition: The implicit association test. *J. Pers. Soc. Psychol.* **74**, 1464–1480 (1998).
 22. Payne, B. K., Cheng, C. M., Govorun, O. & Stewart, B. D. An inkblot for attitudes: Affect misattribution as implicit measurement. *J. Pers. Soc. Psychol.* **89**, 277–293 (2005).
 23. Fazio, R. H., Jackson, J. R., Dunton, B. C. & Williams, C. J. Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *J. Pers. Soc. Psychol.* **69**, 1013–1027 (1995).
 24. Ye, Y. & Gawronski, B. Validating the semantic misattribution procedure as an implicit measure of gender stereotyping. *Eur. J. Soc. Psychol.* **48**, 348–364 (2018).
 25. Greenwald, A. G. & Farnham, S. D. Using the Implicit Association Test to measure self-esteem and self-concept. *J. Pers. Soc. Psychol.* **79**, 1022–1038 (2000).
 26. Slabbinck, H., De Houwer, J. & Van Kenhove, P. A pictorial Attitude IAT as a Measure of Implicit Motives. *Eur. J. Personal.* **25**, 76–86 (2011).
 27. Marvel, J. D. & Resh, W. D. An Unconscious Drive to Help Others? Using the Implicit Association Test to Measure Prosocial Motivation. *Int. Public Manag. J.* **22**, 29–70 (2019).
 28. Seewald, A., Teige-Mocigemba, S. & Rief, W. Outcome Expectations in Psychotherapy: Validation of the Therapy Single Category Implicit Association Test (Therapy SC-IAT). *Cogn. Ther. Res.* **47**, 894–908 (2023).
 29. Dentale, F., Vecchione, M., Gebauer, J. E. & Barbaranelli, C. Measuring automatic value orientations: The Achievement–Benevolence Implicit Association Test. *Br. J. Soc. Psychol.* **57**, 210–229 (2018).
 30. Randall, J. R., Rowe, B. H., Dong, K. A., Nock, M. K. & Colman, I. Assessment of self-harm risk using implicit thoughts. *Psychol. Assess.* **25**, 714–721 (2013).
 31. Tucker, R. P., Wingate, L. R., Burkley, M. & Wells, T. T. Implicit Association with Suicide as Measured by the Suicide Affect Misattribution Procedure (S-AMP) Predicts Suicide Ideation. *Suicide Life. Threat. Behav.* **48**, 720–731 (2018).
 32. Cummins, J. & De Houwer, J. Are Relational Implicit Measures Sensitive to Relational Information? *Collabra Psychol.* **8**, 38621 (2022).
 33. Cummins, J. & De Houwer, J. The Shape of Belief: Developing a Mousetracking-Based Relational Implicit Measure. *Soc. Psychol. Personal. Sci.* **12**, 1517–1526 (2021).
 34. Hansford, T. G., Intawan, C. & Nicholson, S. P. Snap Judgment: Implicit Perceptions of a (Political) Court. *Polit. Behav.* **40**, 127–147 (2018).

35. Schreger, C. & Kimble, M. Assessing civilian perceptions of combat veterans: An IAT study. *Psychol. Trauma Theory Res. Pract. Policy* **9**, 12–18 (2017).
36. Sun, Y., Li, Y., Cai, B. & Li, Q. Comparing the explicit and implicit attitudes of energy stakeholders and the public towards carbon capture and storage. *J. Clean. Prod.* **254**, 120051 (2020).
37. Coleman, J. A., Ingram, K. M., Bays, A., Joy-Gaba, J. A. & Boone, E. L. Disability and Assistance Dog Implicit Association Test: A novel IAT. *Rehabil. Psychol.* **60**, 17–26 (2015).
38. Houben, K., Nosek, B. A. & Wiers, R. W. Seeing the forest through the trees: A comparison of different IAT variants measuring implicit alcohol associations. *Drug Alcohol Depend.* **106**, 204–211 (2010).
39. Corneille, O. & Hütter, M. Implicit? What Do You Mean? A Comprehensive Review of the Delusive Implicitness Construct in Attitude Research. *Personal. Soc. Psychol. Rev.* **24**, 212–232 (2020).
40. Gawronski, B., & De Houwer, J. Implicit measures in social and personality psychology. in *Handbook of research methods in social and personality psychology* (eds H. T. Reis & C. M. Judd) 283–310 (Cambridge University Press, 2014).
41. Kurdi, B., Sanchez, A., Dasgupta, N. & Banaji, M. R. (when) do counterattitudinal exemplars shift implicit racial evaluations? Replications and extensions of Dasgupta and Greenwald (2001). *J. Pers. Soc. Psychol.* **126**, 543–565 (2024).
42. Dasgupta, N. & Greenwald, A. G. On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *J. Pers. Soc. Psychol.* **81**, 800–814 (2001).
43. Schwarz, N. & Bohner, G. The construction of attitudes. in *Blackwell Handbook of Social Psychology: Intraindividual Processes* (eds. Tesser, A. & Schwarz, N.) 436–457 (Wiley, 2001). doi:10.1002/9780470998519.ch20.
44. Schwarz, N. & Strack, F. Context Effects in Attitude Surveys: Applying Cognitive Theory to Social Research. *Eur. Rev. Soc. Psychol.* **2**, 31–50 (1991).
45. Tourangeau, R., Rips, L. J. & Rasinski, K. *The Psychology of Survey Response*. (Cambridge University Press, 2000). doi:10.1017/CBO9780511819322.
46. Strack, F., Martin, L. L. & Schwarz, N. Priming and communication: Social determinants of information use in judgments of life satisfaction. *Eur. J. Soc. Psychol.* **18**, 429–442 (1988).
47. Blair, I. V. The Malleability of Automatic Stereotypes and Prejudice. *Personal. Soc. Psychol. Rev.* **6**, 242–261 (2002).
48. Forscher, P. S. *et al.* A meta-analysis of procedures to change implicit measures. *J. Pers. Soc. Psychol.* **117**, 522–559 (2019).
49. Gawronski, B. & Sritharan, R. Formation, change, and contextualization of mental associations: Determinants and principles of variations in implicit measures. in *Handbook of implicit social cognition: Measurement, theory, and applications*. (eds. Gawronski, B. & Payne, B. K.) 216–240 (New York, NY, 2010).
50. Gawronski, B., Brownstein, M. & Madva, A. How should we think about implicit measures and their empirical “anomalies”? *WIREs Cogn. Sci.* **13**, e1590 (2022).
51. Lelkes, Y., Krosnick, J. A., Marx, D. M., Judd, C. M. & Park, B. Complete anonymity compromises the accuracy of self-reports. *J. Exp. Soc. Psychol.* **48**, 1291–1299 (2012).
52. Berry, B. A. Experimenter Characteristics, Social Desirability, and the Implicit Association Test. *Psi Chi J. Psychol. Res.* **20**, 247–257 (2015).
53. Lowery, B. S., Hardin, C. D. & Sinclair, S. Social influence effects on automatic racial prejudice. *J. Pers. Soc. Psychol.* **81**, 842–855 (2001).
54. Boysen, G. A., Vogel, D. L. & Madon, S. A public versus private administration of the implicit association test. *Eur. J. Soc. Psychol.* **36**, 845–856 (2006).
55. Nosek, B. A. Moderators of the Relationship Between Implicit and Explicit Evaluation. *J. Exp. Psychol. Gen.* **134**, 565–584 (2005).
56. Schmidt, K., Buchanan, E. M. & Hall, B. F. Registered report: Moderators of the relationship between implicit and explicit measures of evaluation and identification. *J. Exp. Soc. Psychol.* **100**, 104280 (2022).
57. Cameron, C. D., Brown-Iannuzzi, J. L. & Payne, B. K. Sequential Priming Measures of Implicit Social Cognition: A Meta-Analysis of Associations With Behavior and Explicit Attitudes. *Personal. Soc. Psychol. Rev.* **16**, 330–350 (2012).
58. Hofmann, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. *Pers. Soc. Psychol. Bull.* **31**, 1369–1385 (2005).
59. Turner, R. N. & Crisp, R. J. Imagining intergroup contact reduces implicit prejudice. *Br. J. Soc. Psychol.* **49**, 129–142 (2010).
60. Zaltman, G. *How Customers Think: Essential Insights into the Mind of the Market*. (Harvard Business School Press, Boston, Mass, 2007).
61. Crano, W. D. & Prislin, R. *Attitudes and Attitude Change*. (Psychology Press, 2011). doi:10.4324/9780203838068.
62. Corneille, O. & Lush, P. Sixty Years After

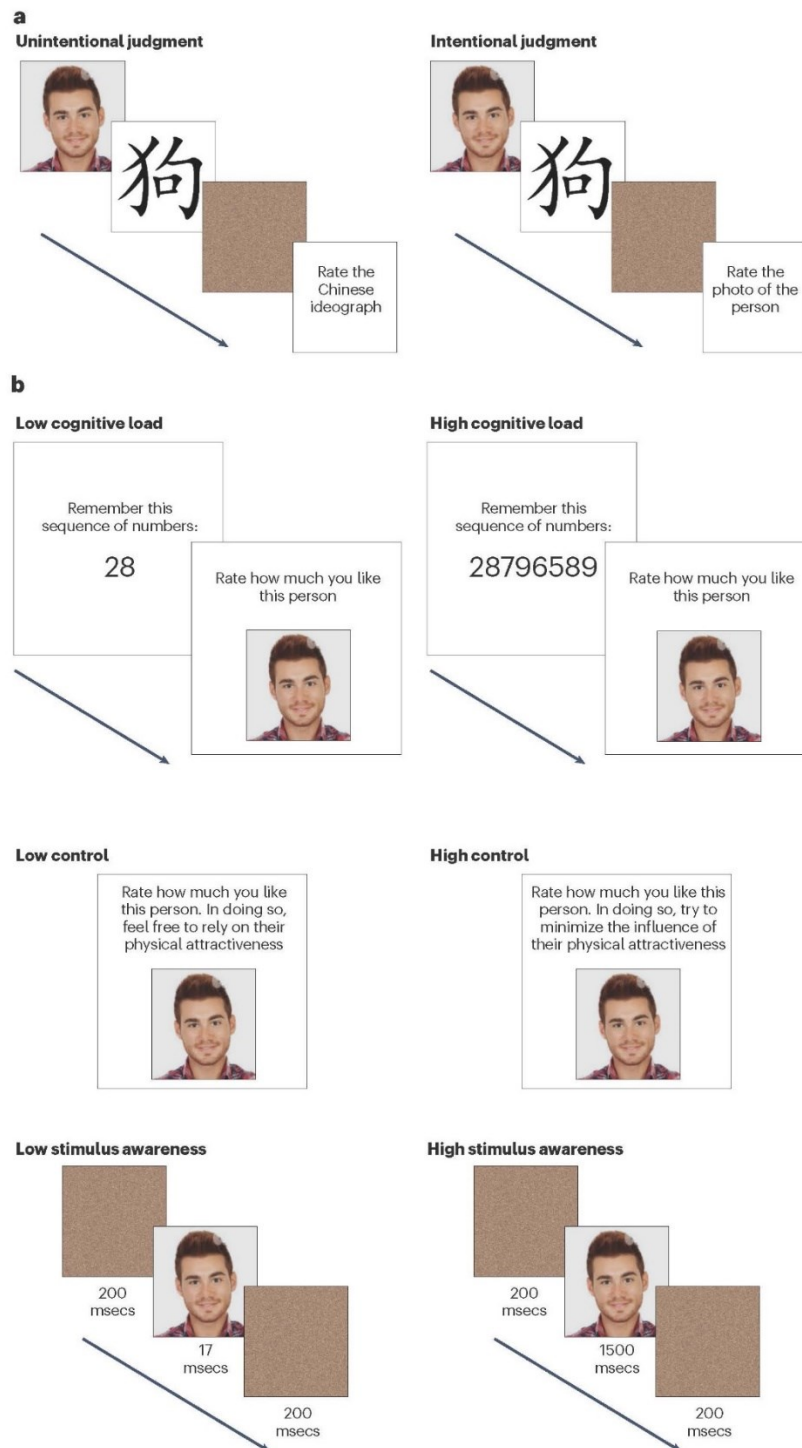
- Orne's *American Psychologist* Article: A Conceptual Framework for Subjective Experiences Elicited by Demand Characteristics. *Personal. Soc. Psychol. Rev.* **27**, 83–101 (2023).
63. Röhner, J., Thoss, P. & Schütz, A. Lying on the Dissection Table: Anatomizing Faked Responses. *Behav. Res. Methods* **54**, 2878–2904 (2022).
64. Steffens, M. C. Is the Implicit Association Test Immune to Faking? *Exp. Psychol.* **51**, 165–179 (2004).
65. Teige-Mocigemba, S., Penzl, B., Becker, M., Henn, L. & Klauer, K. C. Controlling the “uncontrollable”: Faking effects on the affect misattribution procedure. *Cogn. Emot.* **30**, 1470–1484 (2016).
66. Teige-Mocigemba, S. & Klauer, K. C. On the controllability of evaluative-priming effects: Some limits that are none. *Cogn. Emot.* **27**, 632–657 (2013).
67. Fiedler, K. & Bluemke, M. Faking the IAT: Aided and Unaided Response Control on the Implicit Association Tests. *Basic Appl. Soc. Psychol.* **27**, 307–316 (2005).
68. Van Dessel, P. & De Houwer, J. Hypnotic Suggestions Can Induce Rapid Change in Implicit Attitudes. *Psychol. Sci.* **30**, 1362–1370 (2019).
69. Blair, I. V., Ma, J. E. & Lenton, A. P. Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *J. Pers. Soc. Psychol.* **81**, 828–841 (2001).
70. Dhammapeera, P., Hu, X. & Bergström, Z. M. Imagining a false alibi impairs concealed memory detection with the autobiographical Implicit Association Test. *J. Exp. Psychol. Appl.* **26**, 266–282 (2020).
71. Dienes, Z. & Lush, P. The Role of Phenomenological Control in Experience. *Curr. Dir. Psychol. Sci.* **32**, 145–151 (2023).
72. Banaji, M. R. & Greenwald, A. G. *Blindspot: Hidden Biases of Good People*. (Delacorte Press., New York: Bantam, 2013).
73. Kurdi, B. & Mandelbaum, E. The case against implicit bias fatalism. *Nat. Rev. Psychol.* **2**, 656–657 (2023).
74. Morehouse, K. N. & Banaji, M. R. The Science of Implicit Race Bias: Evidence from the Implicit Association Test. *Daedalus* **153**, 21–50 (2024).
75. De Houwer, J., Teige-Mocigemba, S., Spruyt, A. & Moors, A. Implicit measures: A normative analysis and review. *Psychol. Bull.* **135**, 347–368 (2009).
76. Gawronski, B., Hofmann, W. & Wilbur, C. J. Are “implicit” attitudes unconscious? *Conscious. Cogn.* **15**, 485–499 (2006).
77. Gawronski, B. & Corneille, O. Unawareness of Attitudes, Their Environmental Causes, and Their Behavioral Effects. *Annu. Rev. Psychol.* (2024) doi:10.1146/annurev-psych-051324-031037.
78. Hahn, A., Judd, C. M., Hirsh, H. K. & Blair, I. V. Awareness of implicit attitudes. *J. Exp. Psychol. Gen.* **143**, 1369–1392 (2014).
79. Morris, A. & Kurdi, B. Awareness of implicit attitudes: Large-scale investigations of mechanism and scope. *J. Exp. Psychol. Gen.* **152**, 3311–3343 (2023).
80. Rahmani Azad, Z., Goedderz, A. & Hahn, A. Self-Awareness and Stereotypes: Accurate Prediction of Implicit Gender Stereotyping. *Pers. Soc. Psychol. Bull.* **49**, 1695–1708 (2023).
81. Hahn, A. & Gawronski, B. Facing one's implicit biases: From awareness to acknowledgment. *J. Pers. Soc. Psychol.* **116**, 769–794 (2019).
82. Hahn, A. & Gawronski, B. Do implicit evaluations reflect unconscious attitudes? *Behav. Brain Sci.* **37**, 28–29 (2014).
83. Krickel, B. Are the states underlying implicit biases unconscious? – A Neo-Freudian answer. *Philos. Psychol.* **31**, 1007–1026 (2018).
84. Banaji, M. R. A vehicle for large-scale education about the human mind. in *How is the internet changing the way you think?* 392–395 (Harper Collins, New York, 2011).
85. Wolsiefer, K., Westfall, J. & Judd, C. M. Modeling stimulus variation in three common implicit attitude tasks. *Behav. Res. Methods* **49**, 1193–1209 (2017).
86. Gawronski, B., Ledgerwood, A. & Eastwick, P. W. Implicit Bias ≠ Bias on Implicit Measures. *Psychol. Inq.* **33**, 139–155 (2022).
87. Greenwald, A. G. *et al.* Implicit-Bias Remedies: Treating Discriminatory Bias as a Public-Health Problem. *Psychol. Sci. Public Interest* **23**, 7–40 (2022).
88. Greenwald, A. G. & Lai, C. K. Implicit Social Cognition. *Annu. Rev. Psychol.* **71**, 419–445 (2020).
89. Bartels, J. M. & Schoenrade, P. The Implicit Association Test in Introductory Psychology Textbooks: Blind Spot for Controversy. *Psychol. Learn. Teach.* **21**, 113–125 (2022).
90. Mitchell, G. & Tetlock, P. E. Stretching the Limits of Science: Was the Implicit-Racism Debate a “Bridge Too Far” for Social Psychology? in *The Cambridge Handbook of Implicit Bias and Racism* (eds. Krosnick, J. A., Stark, T. H., & Scott, A. L.) (Cambridge University Press., 2024).
91. Balas, R. & Gawronski, B. On the intentional control of conditioned evaluative responses. *Learn. Motiv.* **43**, 89–98 (2012).
92. Ranganath, K. A., Smith, C. T. & Nosek, B. A. Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *J. Exp. Soc. Psychol.* **44**, 386–396 (2008).

93. Stahl, C., Bena, J., Aust, F., Mierop, A. & Corneille, O. A conditional judgment procedure for probing evaluative conditioning effects in the absence of feelings of remembering. *Behav. Res. Methods* **56**, 1140–1163 (2023).
94. Hütter, M. & Klauer, K. C. Applying processing trees in social psychology. *Eur. Rev. Soc. Psychol.* **27**, 116–159 (2016).
95. Payne, B. K. & Bishara, A. J. An integrative review of process dissociation and related models in social cognition. *Eur. Rev. Soc. Psychol.* **20**, 272–314 (2009).
96. Béna, J., Lacassagne, D. & Corneille, O. Do uncontrolled processes contribute to evaluative learning? Insights from a new two-US process dissociation procedure and ambivalence measures. *Q. J. Exp. Psychol.* 17470218241248138 (2024) doi:10.1177/17470218241248138.
97. Hütter, M. & Sweldens, S. Dissociating controllable and uncontrollable effects of affective stimuli on attitudes and consumption. *J. Consum. Res.* **45**, 320–349 (2018).
98. Corneille, O., Mierop, A., Stahl, C. & Hütter, M. Evidence suggestive of uncontrollable attitude acquisition replicates in an instructions-based evaluative conditioning paradigm: Implications for associative attitude acquisition. *J. Exp. Soc. Psychol.* **85**, 103841 (2019).
99. Hansen, J. & Wänke, M. Liking What's Familiar: The Importance of Unconscious Familiarity in the Mere-Exposure Effect. *Soc. Cogn.* **27**, 161–182 (2009).
100. Hütter, M., Sweldens, S., Stahl, C., Unkelbach, C. & Klauer, K. C. Dissociating contingency awareness and conditioned attitudes: Evidence of contingency-unaware evaluative conditioning. *J. Exp. Psychol. Gen.* **141**, 539–557 (2012).
101. Jacoby, L. L. A process dissociation framework: Separating automatic from intentional uses of memory. *J. Mem. Lang.* **30**, 513–541 (1991).
102. Mierop, A., Hütter, M., Stahl, C. & Corneille, O. Does attitude acquisition in evaluative conditioning without explicit CS-US memory reflect implicit misattribution of affect? *Cogn. Emot.* **33**, 173–184 (2019).
103. Unkelbach, C. & Stahl, C. A multinomial modeling approach to dissociate different components of the truth effect. *Conscious. Cogn.* **18**, 22–38 (2009).
104. Strack, F. & Deutsch, R. Reflective and Impulsive Determinants of Social Behavior. *Personal. Soc. Psychol. Rev.* **8**, 220–247 (2004).
105. De Houwer, J., Van Dessel, P. & Moran, T. Attitudes beyond associations: On the role of propositional representations in stimulus evaluation. in *Advances in Experimental Social Psychology* vol. 61 127–183 (Elsevier, 2020).
106. Kurdi, B., Morehouse, K. N. & Dunham, Y. How do explicit and implicit evaluations shift? A preregistered meta-analysis of the effects of co-occurrence and relational information. *J. Pers. Soc. Psychol.* **124**, 1174–1202 (2023).
107. Heycke, T. & Gawronski, B. Co-occurrence and relational information in evaluative learning: A multinomial modeling approach. *J. Exp. Psychol. Gen.* **149**, 104–124 (2020).
108. Kukken, N., Hütter, M. & Holland, R. W. Are there two independent evaluative conditioning effects in relational paradigms? Dissociating the effects of CS-US pairings and their meaning. *Cogn. Emot.* **34**, 170–187 (2020).
109. Béna, J., Mauclet, A. & Corneille, O. Does co-occurrence information influence evaluations beyond relational meaning? An investigation using self-reported and mouse-tracking measures of attitudinal ambivalence. *J. Exp. Psychol. Gen.* **152**, 968–992 (2023).
110. Hu, X., Gawronski, B. & Balas, R. Propositional Versus Dual-Process Accounts of Evaluative Conditioning: I. The Effects of Co-Occurrence and Relational Information on Implicit and Explicit Evaluations. *Pers. Soc. Psychol. Bull.* **43**, 17–32 (2017).
111. Rydell, R. J. & McConnell, A. R. Understanding implicit and explicit attitude change: A systems of reasoning analysis. *J. Pers. Soc. Psychol.* **91**, 995–1008 (2006).
112. Wilson, T. D., Lindsey, S. & Schooler, T. Y. A model of dual attitudes. *Psychol. Rev.* **107**, 101–126 (2000).
113. Rudman, L. A. Sources of Implicit Attitudes. *Curr. Dir. Psychol. Sci.* **13**, 79–82 (2004).
114. Rydell, R. J., McConnell, A. R., Strain, L. M., Claypool, H. M. & Hugenberg, K. Implicit and explicit attitudes respond differently to increasing amounts of counterattitudinal information. *Eur. J. Soc. Psychol.* **37**, 867–878 (2007).
115. Cone, J., Mann, T. C. & Ferguson, M. J. Changing Our Implicit Minds: How, When, and Why Implicit Evaluations Can Be Rapidly Revised. in *Advances in Experimental Social Psychology* vol. 56 131–199 (Elsevier, 2017).
116. Kurdi, B., Mann, T. C. & Ferguson, M. J. Persuading the Implicit Mind: Changing Negative Implicit Evaluations With an 8-Minute Podcast. *Soc. Psychol. Personal. Sci.* **13**, 688–697 (2022).
117. Van Dessel, P., Ye, Y. & De Houwer, J. Changing Deep-Rooted Implicit Evaluation in the Blink of an Eye: Negative Verbal Information Shifts Automatic Liking of Gandhi. *Soc. Psychol. Personal. Sci.* **10**, 266–273 (2019).
118. Lai, C. K. *et al.* Reducing implicit racial

- preferences: II. Intervention effectiveness across time. *J. Exp. Psychol. Gen.* **145**, 1001–1016 (2016).
119. Charlesworth, T. E. S. & Banaji, M. R. Patterns of implicit and explicit attitudes II. Long-term change and stability, regardless of group membership. *Am. Psychol.* **76**, 851–869 (2021).
 120. Charlesworth, T. E. S. & Banaji, M. R. Patterns of Implicit and Explicit Attitudes: IV. Change and Stability From 2007 to 2020. *Psychol. Sci.* **33**, 1347–1371 (2022).
 121. Charlesworth, T. E. S. & Banaji, M. R. Patterns of Implicit and Explicit Stereotypes III: Long-Term Change in Gender Stereotypes. *Soc. Psychol. Personal. Sci.* **13**, 14–26 (2022).
 122. Payne, B. K., Vuletich, H. A. & Lundberg, K. B. The Bias of Crowds: How Implicit Bias Bridges Personal and Systemic Prejudice. *Psychol. Inq.* **28**, 233–248 (2017).
 123. Payne, B. K. & Hannay, J. W. Implicit bias reflects systemic racism. *Trends Cogn. Sci.* **25**, 927–936 (2021).
 124. Ekstrom, P. D., Le Forestier, J. M. & Lai, C. K. Racial Demographics Explain the Link Between Racial Disparities in Traffic Stops and County-Level Racial Attitudes. *Psychol. Sci.* **33**, 497–509 (2022).
 125. Hehman, E., Flake, J. K. & Calanchini, J. Disproportionate Use of Lethal Force in Policing Is Associated With Regional Racial Biases of Residents. *Soc. Psychol. Personal. Sci.* **9**, 393–401 (2018).
 126. Leitner, J. B., Hehman, E., Ayduk, O. & Mendoza-Denton, R. Racial bias is associated with ingroup death rate for Blacks and Whites: Insights from Project Implicit. *Soc. Sci. Med.* **170**, 220–227 (2016).
 127. Galvan, M. J. & Payne, B. K. Implicit Bias as a Cognitive Manifestation of Systemic Racism. *Daedalus* **153**, 106–122 (2024).
 128. Calanchini, J. *et al.* Regional intergroup bias. in *Advances in Experimental Social Psychology* vol. 66 281–337 (Elsevier, 2022).
 129. Orchard, J. & Price, J. County-level racial prejudice and the black-white gap in infant health outcomes. *Soc. Sci. Med.* **181**, 191–198 (2017).
 130. Stelter, M., Essien, I., Sander, C. & Degner, J. Racial Bias in Police Traffic Stops: White Residents' County-Level Prejudice and Stereotypes Are Related to Disproportionate Stopping of Black Drivers. *Psychol. Sci.* **33**, 483–496 (2022).
 131. Swim, J. K., Aikin, K. J., Hall, W. S. & Hunter, B. A. Sexism and racism: Old-fashioned and modern prejudices. *J. Pers. Soc. Psychol.* **68**, 199–214 (1995).
 132. Henry, P. J. & Sears, D. O. The Symbolic Racism 2000 Scale. *Polit. Psychol.* **23**, 253–283 (2002).
 133. McConahay, J. B. Modern racism, ambivalence, and the Modern Racism Scale. in *Prejudice, discrimination, and racism* (eds. Dovidio, J.F. & Gaertner, S. L.) 91–125 (Academic Press, 1986).
 134. Gawronski, B., Morrison, M., Phillips, C. E. & Galdi, S. Temporal Stability of Implicit and Explicit Measures: A Longitudinal Analysis. *Pers. Soc. Psychol. Bull.* **43**, 300–312 (2017).
 135. Koppehele-Gossel, J., Hoffmann, L., Banse, R. & Gawronski, B. Evaluative priming as an implicit measure of evaluation: An examination of outlier-treatments for evaluative priming scores. *J. Exp. Soc. Psychol.* **87**, 103905 (2020).
 136. LeBel, E. P. & Paunonen, S. V. Sexy But Often Unreliable: The Impact of Unreliability on the Replicability of Experimental Findings With Implicit Measures. *Pers. Soc. Psychol. Bull.* **37**, 570–583 (2011).
 137. Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L. & Banaji, M. R. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *J. Pers. Soc. Psychol.* **97**, 17–41 (2009).
 138. Kurdi, B. *et al.* Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *Am. Psychol.* **74**, 569–586 (2019).
 139. Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J. & Tetlock, P. E. Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *J. Pers. Soc. Psychol.* **105**, 171–192 (2013).
 140. Ajzen, I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **50**, 179–211 (1991).
 141. Albarracín, D., Johnson, B. T., Fishbein, M. & Muellerleile, P. A. Theories of reasoned action and planned behavior as models of condom use: A meta-analysis. *Psychol. Bull.* **127**, 142–161 (2001).
 142. Lin, N. & Roberts, K. R. Using the theory of planned behavior to predict food safety behavioral intention: A systematic review and meta-analysis. *Int. J. Hosp. Manag.* **90**, 102612 (2020).
 143. McDermott, M. S. *et al.* The Theory of Planned Behaviour and dietary patterns: A systematic review and meta-analysis. *Prev. Med.* **81**, 150–156 (2015).
 144. McEachan, R. R. C., Conner, M., Taylor, N. J. & Lawton, R. J. Prospective prediction of health-related behaviours with the Theory of Planned Behaviour: a meta-analysis. *Health Psychol. Rev.* **5**, 97–144 (2011).
 145. Albarracín, D., Fayaz-Farkhad, B. & Granados Samayoa, J. A. Determinants of behaviour and their efficacy as targets of behavioural change interventions. *Nat. Rev. Psychol.* **3**, 377–392 (2024).
 146. Glasman, L. R. & Albarracín, D. Forming attitudes that predict future behavior: A meta-analysis

- of the attitude-behavior relation. *Psychol. Bull.* **132**, 778–822 (2006).
147. Frieze, M., Hofmann, W. & Schmitt, M. When and why do implicit measures predict behaviour? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *Eur. Rev. Soc. Psychol.* **19**, 285–338 (2008).
 148. Fazio, R. H. Attitudes as Object–Evaluation Associations of Varying Strength. *Soc. Cogn.* **25**, 603–637 (2007).
 149. Perugini, M., Richetin, J., & Zogmaister, C. Prediction of behavior. in *Handbook of implicit social cognition: Measurement, theory, and applications* (eds. Gawronski, B. & Payne, B. K.) 255–277 (The Guilford Press, 2010).
 150. Gawronski, B. Six Lessons for a Cogent Science of Implicit Bias and Its Criticism. *Perspect. Psychol. Sci.* **14**, 574–595 (2019).
 151. Carpenter, T. P., Goedderz, A. & Lai, C. K. Individual Differences in Implicit Bias Can Be Measured Reliably by Administering the Same Implicit Association Test Multiple Times. *Pers. Soc. Psychol. Bull.* **49**, 1363–1378 (2023).
 152. Kaiser, C. & Oswald, A. J. The scientific value of numerical measures of human feelings. *Proc. Natl. Acad. Sci.* **119**, e2210412119 (2022).
 153. Vargas, P. T., Von Hippel, W. & Petty, R. E. Using Partially Structured Attitude Measures to Enhance the Attitude-Behavior Relationship. *Pers. Soc. Psychol. Bull.* **30**, 197–211 (2004).
 154. Bodenhausen, G. V., & Petsko, C. D. Complications in predicting intergroup behavior from implicit biases: One size does not fit all. in *The Cambridge handbook of implicit bias and racism*. (eds. Krosnick, J. A., Stark, T. H., & Scott, A. L.) (Cambridge University Press, 2024).
 155. Payne, B. K., Burkley, M. A. & Stokes, M. B. Why do implicit and explicit attitude tests diverge? The role of structural fit. *J. Pers. Soc. Psychol.* **94**, 16–31 (2008).
 156. Béna, J., Melnikoff, D. E., Mierop, A. & Corneille, O. Revisiting dissociation hypotheses with a structural fit approach: The case of the prepared reflex framework. *J. Exp. Soc. Psychol.* **100**, 104297 (2022).
 157. Zerhouni, O., Bègue, L., Comiran, F. & Wiers, R. W. Controlled and implicit processes in evaluative conditioning on implicit and explicit attitudes toward alcohol and intentions to drink. *Addict. Behav.* **76**, 335–342 (2018).
 158. Béna, J., Lacassagne, D. & Corneille, O. Do uncontrolled processes contribute to evaluative learning? Insights from a new two-US process dissociation procedure and ambivalence measures. *Q. J. Exp. Psychol.* 17470218241248138 (2024) doi:10.1177/17470218241248138.
 159. Gawronski, B. & Bodenhausen, G. V. Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychol. Bull.* **132**, 692–731 (2006).
 160. Calanchini, J. How multinomial processing trees have advanced, and can continue to advance, research using implicit measures. *Soc. Cogn.* **38**, s165–s186 (2020).
 161. De Houwer, J. & Boddez, Y. Bias in Implicit Measures as Instances of Biased Behavior under Suboptimal Conditions in the Laboratory. *Psychol. Inq.* **33**, 173–176 (2022).
 162. Gawronski, B. & Bodenhausen, G. V. Theory evaluation. in *Theory and explanation in social psychology* (eds. Gawronski, B. & Bodenhausen, G. V.) 3–23 (Guilford Press, New York, NY, 2015).
 163. Goedderz, A. & Hahn, A. Awareness and Calibration: The Role of Descriptive Norms and Social Desirability in Accurate IAT Score Predictions of Food Items vs. Social Groups. *Pers. Soc. Psychol. Bull.* 01461672241254447 (2024) doi:10.1177/01461672241254447.
 164. Ratliff, K. A. & Smith, C. T. Implicit Bias as Automatic Behavior. *Psychol. Inq.* **33**, 213–218 (2022).
 165. Kurdi, B. *et al.* Testing the automaticity features of the affect misattribution procedure: The roles of awareness and intentionality. *Behav. Res. Methods* **56**, 3161–3194 (2023).
 166. Schimmack, U. Invalid Claims About the Validity of Implicit Association Tests by Prisoners of the Implicit Social-Cognition Paradigm. *Perspect. Psychol. Sci.* **16**, 435–442 (2021).
 167. Schimmack, U. The Implicit Association Test: A Method in Search of a Construct. *Perspect. Psychol. Sci.* **16**, 396–414 (2021).
 168. Buchner, A., Erdfelder, E. & Vaterrodt-Plünnecke, B. Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *J. Exp. Psychol. Gen.* **124**, 137–160 (1995).
 169. Schmidt, O., Erdfelder, E. & Heck, D. W. How to develop, test, and extend multinomial processing tree models: A tutorial. *Psychol. Methods* (2023) doi:10.1037/met0000561.

Figure 1. Structural-fit approaches for drawing inferences about automaticity. In structural-fit paradigms, responses are compared across two variants of the same task that differ only with regard to the automaticity feature of interest A) Structural fit between an implicit measure (left) and a self-report measure (right) to assess the role of intentionality on evaluations of a prime (face) stimulus. B) Structural fit between two self-report measures to assess resource-dependency (top), control (middle), and stimulus awareness (bottom) on evaluations of a face stimulus.



Box 1. Process Dissociation

Process dissociation enables researchers to quantify the contributions of distinct processes to responses on a given task or measure.^{94,95,160} In its simplest form, process dissociation can be used to analyze responses on tasks or measures with binary outcomes in which two distinct processes can lead to either to the same or different outcomes (for example, correct vs. incorrect judgments; positive vs. negative judgments; or stereotype-congruent vs. stereotype-incongruent judgments). Trials or conditions in which the two processes lead to the same outcome can be described as compatible; trials or conditions in which the two processes lead to different outcomes can be described as incompatible.

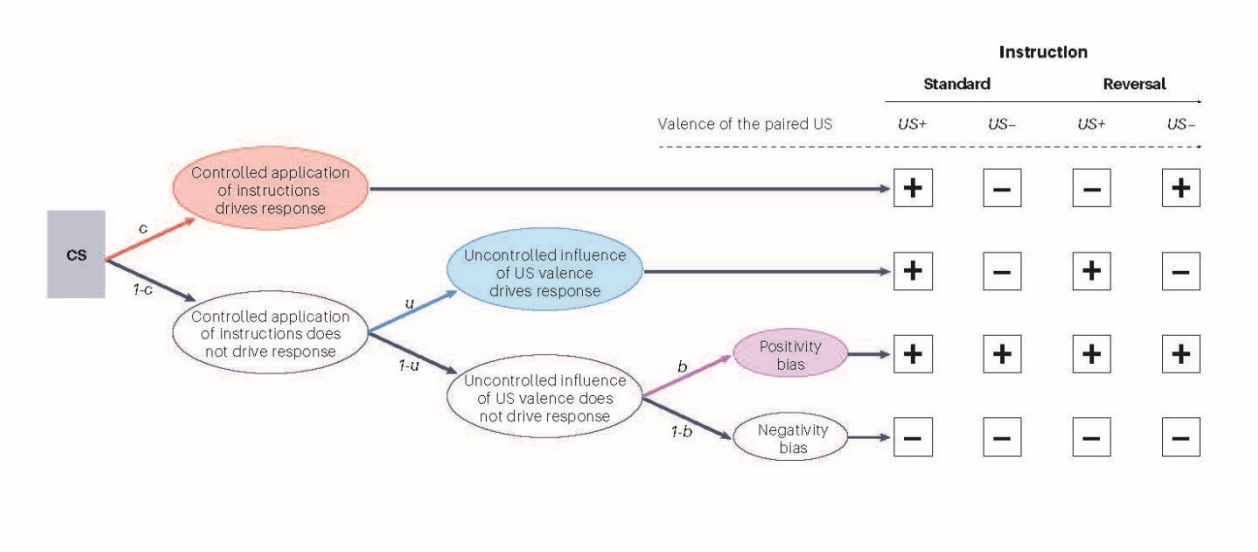
In one version of this paradigm⁹⁷, images of affectively neutral stimuli such as bottled water brand logos (conditioned stimuli) are paired with images of positive valence such as pets or flowers (positive unconditioned stimulus) or negative valence such as garbage or injuries (negative unconditioned stimulus). Depending on the condition, participants are instructed to form an impression of the conditioned stimulus that is consistent with the valence of the unconditioned stimulus (standard condition) or opposite to it (reversal condition). Subsequently, participants are asked to report whether they like or dislike each conditioned stimulus.

In this study design, three parameters can be estimated by comparing the frequencies of positive and negative conditioned stimulus evaluations on compatible trials (where controlled and uncontrolled processes are assumed to have compatible evaluative influences) and incompatible trials (where controlled and uncontrolled processes are assumed to have incompatible evaluative influences). A first parameter, c , reflects participants' tendency to correctly apply the instruction. In this case, responses are driven by controlled processes in the sense that participants show their ability to correctly apply either the standard or reversal instruction (see red arrow in the figure). If controlled processes fail to drive participants' responses ($1-c$), two additional parameters can be estimated. One of these parameters is u , which reflects participants' tendency to form an impression of the conditioned stimulus that is consistent with the valence of the unconditioned stimulus regardless of the specific instruction. In this case, participants' evaluative responses are assumed to be driven by uncontrolled processes (see blue arrow in the figure). The third parameter, b , reflects a general bias to evaluate the conditioned stimulus positively (vs. negatively) when neither c nor u processes are at play (see purple arrow in the figure).

The three parameters can be estimated using a set of mathematical equations. In its simplest form, process dissociation estimates only the controlled (c) and uncontrolled (u) parameters. In this case, the controlled process can be quantified as $c = p(\text{correct} \mid \text{compatible}) - p(\text{incorrect} \mid \text{incompatible})$, and the uncontrolled process can be quantified as $u = p(\text{incorrect} \mid \text{incompatible}) / (1 - c)$. However, an unbiased estimate of these parameters requires additionally estimating a potential decision bias (b).¹⁶⁸ When b is estimated in addition to c and u , the set of equations becomes more complex, requiring the use of multinomial processing tree modeling in which parameters values are estimated via maximum-likelihood statistics. There are useful tutorials that can be used to develop and test multinomial processing tree models.¹⁶⁹

Process dissociation can be seen as an advanced application of structural fit for estimating the independent contributions of various latent cognitive processes to task performance. In the present illustration, controlled and uncontrolled processes can be estimated within a unique task that keeps all procedural factors constant except for the instruction (standard vs. reversal) communicated to the participants.

This illustration of process dissociation emphasizes two important points that are directly relevant to our current analysis. First, self-reports can be used to study automatic processes. Second, self-reports are not entirely immune to uncontrolled influences (provided a good model fit is observed, and the u parameter—or related parameters reflecting the contribution of uncontrolled processes¹⁵⁸—is significantly greater than zero).



Note: CS: conditioned stimuli; US+: positive unconditioned stimulus; US-: negative unconditioned stimulus.