

Methodology and Research Practice

Instruction-based Replication Studies Raise Challenging Questions for Psychological Science

Olivier Corneille¹, Jérémy Béna¹

¹ Psychological Sciences Research Institute, UCLouvain, Louvain-la-Neuve, Belgium

Keywords: Lay Psychological theories, Methods, Demand Effects, Experimental artifacts, Replication, Experimental validity

<https://doi.org/10.1525/collabra.82234>

Collabra: Psychology

Vol. 9, Issue 1, 2023

A variety of psychological effects have been recently replicated in studies where participants merely received information describing experimental tasks, while participants experienced these tasks in studies where these effects were originally established. We argue that these successful instruction-based replication studies raise challenging questions for contemporary psychological research: (1) What does psychological science tell us about effects beyond common knowledge? (2) Does performing the experienced version of the task add to the effect, how much so, and why? (3) Should the effect be considered an experimental demand artifact? Throughout the article, we discuss methodological challenges and solutions associated with these questions. We conclude that instruction-based replication studies offer opportunities for theoretical, methodological, and empirical development in psychological science.

Introduction

Participants taking part in psychological experiments typically experience procedures, the nature of which varies across conditions. For instance, in a typical approach-avoidance experiment, participants may be requested to physically approach or to physically avoid stimuli. At a later stage of the experiment, they may be asked to report how much they like these stimuli. If they report liking better the physically approached than the physically avoided stimuli, an approach-avoidance effect is established (e.g., Cacioppo et al., 1993).

In the present article, we discuss questions that arise when effects originally established in such procedures are successfully replicated in studies where participants are *merely instructed* about the procedures (rather than experiencing them). For example, would participants evaluate more positively stimuli if they are *merely told* that these stimuli will be approached rather than avoided? Recent evidence suggests that effects originally found in experienced procedures often replicate when these procedures are merely described to the participants; that is, in instruction-based replication studies (IBRs, for short).

Successful instruction-based replications of effects originally found in experienced procedures have been reported for a diverse range of phenomena. Merely providing participants with instructions about original procedures was enough to replicate the approach-avoidance effect (i.e., a preference for approached relative to avoided stimuli; Hüt-

ter & Genschow, 2020; Van Dessel et al., 2015, 2016, 2018), the mere exposure effect (i.e., higher liking as a result of repeated exposure; Van Dessel et al., 2017), the evaluative conditioning effect (i.e., a change in liking after pairing neutral stimuli with stimuli of affective valence; Béna, Mauclet, et al., 2022; Corneille et al., 2019; De Houwer, 2006; De Houwer & Hughes, 2016; Gast & De Houwer, 2012; Hütter & De Houwer, 2017; Kurdi & Banaji, 2017, 2019; Mattavelli et al., 2021; Moran et al., 2021), the truth effect (i.e., higher truth judgments as a result of repeated exposure; Mattavelli et al., 2022), and the rubber hand illusion (i.e., an enhanced feeling of ownership reported for a fake hand brushed in synchrony rather than in asynchrony with a participant's real hand concealed to their view; Lush, 2020; Reader, 2022). To illustrate, in a recent instruction-based replication of the mere exposure effect, participants were merely instructed that two words would be presented to them either rarely (e.g., LOKANTA) or frequently (e.g., FEVKANI). Replicating the standard mere exposure effect, being merely instructed about (rather than experiencing) higher stimulus exposure increased stimulus liking (Van Dessel et al., 2017).

IBRs share two distinctive features. First, they provide participants with mere information about the original procedure, but they do not actually implement that procedure, so that participants do not experience the procedure that is described to them. Second, they examine whether effects originally found in their "experienced" counterpart replicate. This definition excludes most "vignette" and verbal

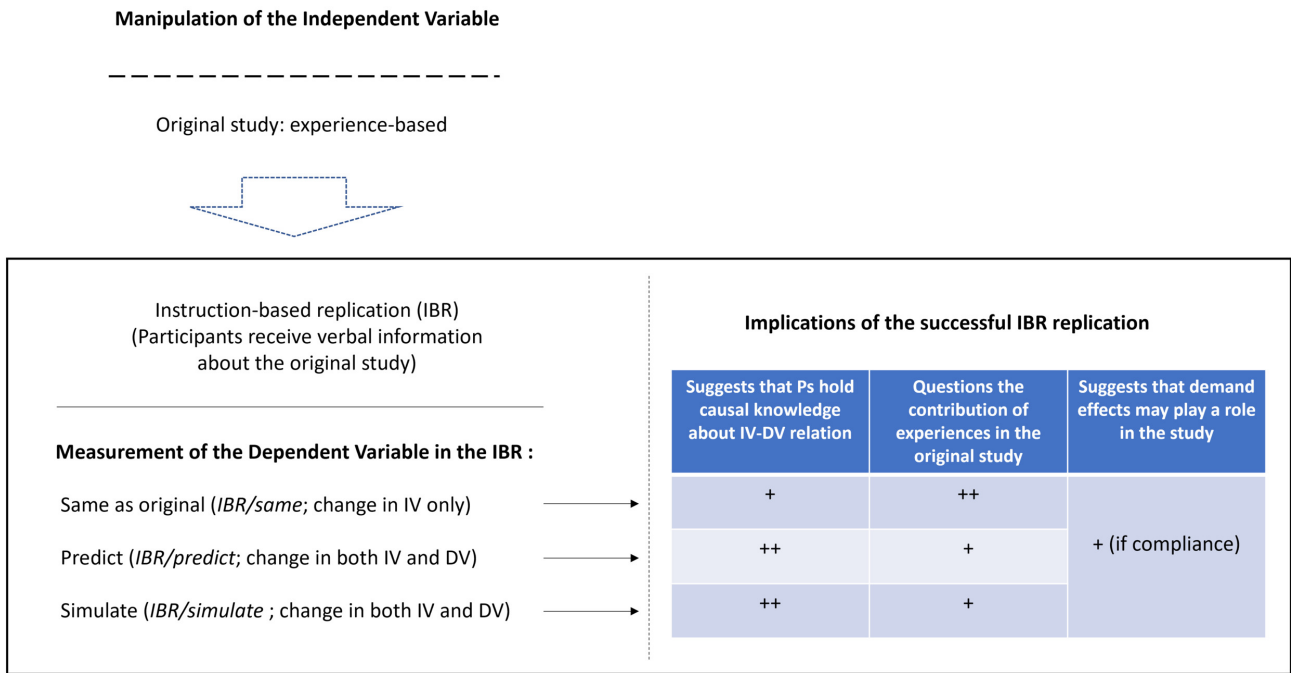


Figure 1. Instruction-based replications and their implications for psychological research.

Note. "+": suggests; "++": strongly suggests

suggestion studies; these studies also produce effects based on mere verbal information, but this information is generally not concerned with procedures used in previous studies.

IBRs are *conceptual* replication studies. The manipulation of the independent variable (IV) is necessarily altered from experiential (e.g., repeated exposure) to verbal manipulations (e.g., *stated* repeated exposure). Provided functional definitions of effects (a change in the dependent variable [DV] due to an IV), altering the IV precludes an exact replication of the effect. Regarding the measure, it can be identical to the original one. However, it may as well depart from it, by either asking participants to predict or to simulate responses on the original DV. Here, we will distinguish three cases: IBRs that rely on the original measure (*IBR/same*; only the IV is changed), IBRs that ask participants to predict responses on the original measure (*IBR/predict*; both the IV and the DV are changed), and IBRs that ask participants to simulate responses on the original measure (*IBR/simulate*; both the IV and the DV are changed; see [Figure 1](#)).¹

An IBR in which both the IV and the DV were changed is the study by Bem (1967). In that study, Bem argued that dissonance effects are driven by cold inferential cognition rather than by a motivation to reduce physiologically aversive states of cognitive dissonance elicited after performing an incongruent behavior. To test this alternative account, Bem (1967) proceeded to a conceptual replication of Fes-

tinger and Carlsmith (1959)'s study in which participants were found to report more positive attitudes about an obnoxious task after agreeing for a small than large reward to persuade another participant that the task was enjoyable. Instead, in Bem's study, participants were merely provided with information about the original study, and they had to predict how a participant in the original study would have completed the measure. Hence, the IV was changed from an experienced to an instruction-based manipulation, and the DV was changed from responses on an attitude scale to interpersonal predictions of responses on the same attitude scale.

Bem (1967) reasoned :*"If this analysis of the findings is correct, then it should be possible to replicate the inverse functional relation between amount of compensation and the final attitude statement by actually letting an outside observer try to infer the attitude of (the subjects) in the original study. Conceptually, this replicates the Festinger-Carlsmith experiment (...)"* (p. 188). It is worth noting that Bem's understanding of "replication" is consistent both with our and contemporary definitions of this notion as *"a study for which any outcome would be considered diagnostic evidence about a claim from prior research"* (Nosek & Errington, 2020, p. 2). This understanding *"reduces emphasis on operational characteristics of the study and increases emphasis on the interpretation of possible outcomes"* (Nosek & Errington, 2020, p. 2). Because the replication was successful, Bem's study challenged the dominant account of dissonance effects: expe-

¹ The measure used in an experience-based study may be a prediction or a simulation measure. In this case, one may conceive of an IBR that provides participants with verbal instructions about that procedure, and asks them to predict or simulate these predictions or simulations. These cases are not discussed here as we are unaware of successful IBRs that proceeded to such test.

riencing dissonance and being motivated to reduce it were not necessary for producing the effect; drawing cold inferences was enough for producing it.

We argue here that successful IBRs raise challenging questions about (1) the extent to which psychological researchers know more about the effect of interest than the layperson, (2) the distinct contribution of task experience posited in prominent psychological theories, and (3) the role of demand effects in the studied phenomenon. We consider these questions in turn. In the concluding sections, we summarize the main questions discussed in this contribution and their theoretical and practical implications, and we call for more IBR research.

1. What does Psychological Science Tell us about Effects beyond Common Knowledge?

When participants can simulate (i.e., *IBR/simulate*) or predict (i.e., *IBR/predict*) effects based on the mere description of a procedure, this suggests that they consciously hold causal knowledge (in the sense of beliefs relating the IV to the DV) about whether and in what direction procedures influence responses. For instance, if participants can predict that stimuli will be liked better after repeated exposure (i.e., the mere exposure effect), there is indication that they consciously hold causal knowledge relating repeated exposures to increased liking. When a successful replication is found in IBRs requesting participants to directly complete the original measure of interest (i.e., *IBR/same*), this also suggests that participants hold causal knowledge relating procedures to responses. In this case, however, there is no indication that this knowledge is consciously held or consciously used by the participant, nor is there indication that it is not (see [Figure 1](#)).

If participants hold knowledge of this sort, then this blurs the boundaries between “naïve” and scientific knowledge. This, in turn, raises disquieting questions. As Houston (1983) pointed out four decades ago in a study finding convergence between lay and scientific psychological knowledge: “*If chemistry, biology, or physics suddenly evaporated, the effects would be immediate and dramatic. It would probably be centuries before the lost information could be regained. But what would happen if psychology suddenly disappeared? Would humanity suffer a severe lesion in stored information, or would the basic principles of psychology be safe in the collective knowledge of the lay population?*” (p. 203; see also Houston, 1985).

How do psychological and naïve knowledge differ then? A first possibility is that, just as researchers, participants hold accurate causal knowledge relating procedures to responses but that, contrary to researchers, participants remain unaware of underlying mechanisms. After all, people have prior knowledge regarding their and others’ behaviors in a wide range of contexts and may use it when it is cued by mere descriptions of situations. By analogy, most people can predict that dropping a stone on the ground will cause it to fall, although explaining exactly why is much more difficult. Psychological researchers, however, do not always agree on underlying processes, or their consensus may only be temporary. For instance, after decades of research on the

evaluative conditioning effect and on the mere exposure effect (two effects successfully replicated in IBRs), there is still no scientific consensus on the mental mechanisms underlying these effects. When scientific consensus is lacking, participants’ knowledge may or may not be in line with the prevailing scientific explanations of the effects (if any).

A second possibility is that researchers, contrary to participants, hold special knowledge about factors that modulate the effect (i.e., the moderators of an effect). For instance, researchers may know under which conditions a larger time interval between exposure and judgment will increase the impact of repeating statements on their perceived truth, while this knowledge may not be part of commonsense. However, one should not posit that such knowledge is necessarily lacking in the participants. Instead, participants’ knowledge of moderation effects should also be tested in IBRs (by providing information about procedures including moderators). To our knowledge, such IBRs are currently lacking.

Finally, researchers may not always be clear about the magnitude or even the direction of causal relations. For instance, a recent study found large disagreements among happiness experts regarding how various strategy policies may contribute to life satisfaction (Buettner et al., 2020). In this case, participants’ and scientific knowledge cannot possibly converge as the latter is ill-defined or contradictory.

Turning to opportunities, successful IBRs suggest that “naïve” knowledge may be a rich source of theoretical development for psychological science. First, one may examine how this knowledge was acquired. For instance, how can participants anticipate that physical approach will increase liking relative to physical avoidance? Do they rely on personal experiences, on observational learning, or do they rely on scientific knowledge they acquired via educational training? Second, successful IBRs suggest that researchers might consider directly asking naïve participants about causal relations involved in psychological phenomena to gain insights from them. Third, unsuccessful IBRs may also inform when, which, and ultimately why participants get this psychological knowledge right or wrong (for a thorough discussion of the interplays between commonsense and scientific psychology, see Kelley, 1992).

This analysis calls for enhanced self-reflection and humility by highlighting that psychological science does not necessarily know much better than the layperson. It also points to unsuspected knowledge-building opportunities that could make use of lay psychological knowledge.

2. Does Performing the Experienced Version of the Task Add to the Effect, How Much so, and Why?

Experience-based procedures involve performing a task and processing verbal information communicated about it in the informed consent and task instructions. Therefore, whenever a successful IBR is achieved, one is left to wonder to what extent experiencing the manipulations contributed to the original effect over and above processing the verbal information provided in the experiment. Identifying this

distinct contribution is theoretically important because prominent psychological theories typically assume that experiences arising during task performance add to the effect, and they often assume that they do so via a distinct set of processes.

For instance, task experiences are assumed to bring a distinct contribution to the following effects: the physical experience of approach/avoidance for embodiment accounts of the approach/avoidance effect (Nuel et al., 2022; Rougier et al., 2018), the experience of repetition-induced fluency for fluency accounts of the mere exposure effect (Leynes & Addante, 2016) and of the truth effect (Unkelbach et al., 2019; Unkelbach & Stahl, 2009), the integration of information coming from different sensory modalities for multisensory accounts of the rubber hand illusion (Ehrsson, 2020; Golaszewski et al., 2021; Kiltner et al., 2015; Tsakiris, 2010), and the experience of spatiotemporal pairings of stimuli for associative or dual-learning accounts of the evaluative conditioning effect (Gawronski et al., 2017).

Several general strategies may be used to assess the distinct contribution of task experiences. We could identify four such strategies. The first two strategies imply making comparisons between instruction-based and experience-based procedures. The last two strategies can be implemented within experience-based procedures only. As we argue below, each strategy comes with significant shortcomings.

Magnitude test: A first strategy has consisted in randomly assigning participants to instruction-based vs. experience-based versions of an experiment and comparing the magnitude of the effect established in these procedures (Forster et al., 2022; Kasran et al., 2022; Kurdi & Banaji, 2017; Lush, 2020; Mattavelli et al., 2022; Moran et al., 2021; Rougier et al., 2021; Van Dessel, De Houwer, et al., 2020). If a larger effect is observed in the experience-based procedure, this may suggest an additional contribution of performing the task (and, possibly, the existence of distinct processes underlying this additional contribution). A refined variation of this strategy consists in modeling and comparing estimates of the processes assumed to contribute to responses produced on the measure of interest (for instance, using multinomial processing tree modeling; Corneille et al., 2019; Hütter & De Houwer, 2017; Smith et al., 2020).

This comparative strategy, however, should guarantee that information is held constant in the experienced vs. instruction-based versions of the procedures, and that these procedures vary only in that one is performed and the other is merely described. This structural fit (i.e., introducing a difference between tasks only on the dimension of interest) is often difficult to achieve. In particular, researchers running an IBR face the question of how detailed the description of the experienced procedure (including stimuli materials) should be. Structural fit in IBRs largely varies, ranging from IBRs that closely (e.g., Bem, 1967; Forster et al., 2022; Lush, 2020) to loosely match their experience-based counterpart.

In addition, outcomes may vary as a function of the procedural operationalization selected. For instance, instruction-based approach-avoidance effects have been robustly

replicated in several procedures, except for one specific procedure crossing a new approach-avoidance manipulation with a new indirect measure of approach-avoidance effects (Rougier et al., 2021), and they have not been found to change evaluations about familiar social groups (Van Dessel, De Houwer, et al., 2020).

Functional test: A second strategy is to test functional similarities and differences between experience-based and instruction-based procedures. Based on the logic of functional dissociations, the rationale is the following: if effects observed in these procedures are driven by the same process(es), then they should be similarly moderated by the same (experimental or individual) factors (e.g., Hu et al., 2017; Kurdi & Banaji, 2019; Moran et al., 2021; Olsson & Phelps, 2007). For instance, both experience-based and instruction-based fear conditioning influence physiological reactions to a conditioned stimulus when it is presented supraliminally. However, only experience-based fear conditioning may influence physiological reactions to a subliminal stimulus (Olsson & Phelps, 2007). This dissociation suggests that distinct processes may be involved in verbal fear conditioning.

Unfortunately, this second comparative strategy faces the same limitations as the previous one, including the difficulty of equating verbal information across conditions (Wisniewski et al., 2022). It is also more challenging because appropriate tests of the higher-order interaction (i.e., whether the moderation of the effect is moderated by the experienced vs. instruction-based nature of the procedure) are likely to require large sample sizes to have enough statistical power (depending on the size of the higher-order interaction effect).

Functional tests may involve comparing the effects of instruction-based vs. experience-based procedures across measures (e.g., for more or less controllable measures of liking; e.g., Hu et al., 2017; Kurdi & Banaji, 2019; Moran et al., 2021). Here, the test of interest is whether differences observed across measures are constant across experience-based and instruction-based procedures. A sound interpretation of such studies, however, requires a structural fit in the measures in addition to a structural fit in the procedures: the two measures should be similar, except for the theoretical factor of interest (Payne et al., 2008; Van Dessel, Cone, et al., 2020). For instance, if the controllability of the measure is the theoretical factor of interest for interpreting a difference between experience-based and instruction-based procedures, then the two measures should differ in their degree of controllability only.

Unfortunately, structural fit can be difficult to achieve, and it is rarely implemented in dissociation tests comparing effects across measures outside, for instance, memory research (e.g., Merikle & Reingold, 1991; Schacter et al., 1989). When comparing performance between structurally equated tasks, conclusions may diverge from conclusions drawn from comparisons between tasks that vary on more than one dimension (Béna, Melnikoff, et al., 2022). Of note too, finding no difference in outcomes between unfitted tasks for experience-based vs. instruction-based manipulations would not support the operation of identical

processes. This is because these procedures may result in the same outcome via different processes. When outcomes on unfitted measures are compared, the interpretational ambiguity can only be worsened, as it becomes virtually impossible to isolate contributing factors.

Deterioration test: Related to the functional test approach, a third strategy assumes that one can distinguish between two categories of processes, one of which is more automatic than the other. For instance, inferential processes drawing from symbolic information (e.g., drawing inferences about the meaning of approach and avoidance) have sometimes been conceptualized as non-automatic; in contrast, embodied/sensory/affective/impulsive/experiential/associative processes (e.g., activating sensory-motor representations while physically approaching/avoiding a stimulus) are generally conceptualized as automatic. If that is the case, (part of) the effect found in an experience-based procedure should remain under conditions thought to compromise the inferential processes but not the other one(s). For instance, when “automatic” refers to efficiency (i.e., relative independence from cognitive resources), the attentional burden of a concurrent task should disrupt “non-automatic” inferences but leave the other “automatic” process(es) unaffected.

Although intuitive, this strategy is problematic in at least two regards. First, processing may be compromised to such a degree that the original experience is fundamentally changed. Second, this strategy rests on the questionable assumption that inferences cannot be automatic (in our current example, that they cannot be efficient; see, e.g., De Houwer, 2018b; Reyna & Brainerd, 1995; Shimizu et al., 2017).

Reversal test: A fourth strategy is to experimentally induce opposite causal inferences within an experience-based procedure. If the effect fails to completely reverse when reversing inferences, this suggests that not just inferences are driving the effect. However, this strategy requires successfully reversing participants’ causal inferences. This may not always be possible. For instance, a recent study could not persuade participants that asynchronous brushing would produce a larger rubber hand illusion than synchronous brushing (Lush et al., 2020). Likewise, researchers recently failed to successfully manipulate expectancies about the efficiency of a clinical treatment (i.e., *Eye-Movement Desensitization Reprocessing therapy*; Mertens et al., 2021). This, by the way, illustrates that sensitive and diagnostic manipulation checks can be important in experimental research (Fiedler et al., 2021). Had the manipulation checks not been collected, the absence of complete reversal may have been wrongly interpreted as evidence for the contribution of non-inferential processes.

It is important to stress that successful IBRs do not imply that “experiential” theories are wrong. However, they challenge these theories to bring evidence for the distinct contribution they posit. To make things even more challenging, this contribution may come in two forms: an independent contribution or an interactive one. We believe (1) that these contributions may largely vary across phenomena, (2) that, for each concerned phenomenon, it would be important to

know better how large these contributions are and whether they are driven by distinct processes, and (3) that one reason we do not know better yet is that answering these questions is methodologically challenging (see above). This calls for more IBR research, and for methodological (e.g., designing structurally fitted procedures) and modeling (testable formal models such as multinomial processing tree models) advances.

Figure 1 summarizes the implications of the three categories of IBRs for our current question. In the case of *IBR/same* studies, the challenge to experiential theories is high because the DV is kept constant. In the case of *IBR/predict* and *IBR/simulate* studies, the challenge is weaker because the DV is altered. It is possible, for instance, that simulations or predictions involve processes that are not engaged in the responses produced in the original procedures. In all three cases, there is also a challenge insofar as the methodological approach is sound. As we discussed here, IBRs can vary in this regard, and we suspect that any method will come with its own set of limitations. Using several methods to build on their strengths and overcome their specific shortcomings would help shed light on the questions we just discussed.

3. Should the Effect be Considered an Experimental Demand Artifact?

Successful IBRs suggest that participants can draw causal inferences relating procedural information to responses, and that these inferences may contribute to effects. Of note, participants may draw broader inferences in a testing situation, including inferences about their role in the situation. In particular, they may be motivated to produce responses consistent with their causal inferences (i.e., to produce the expected response), because doing so feels nice to the experimenter, thoughtful, sensible, easy, interesting, or pleasant in the testing situation. Alternatively, participants may want to affirm their sense of autonomy by producing responses opposite to their causal inferences (e.g., as for reactant participants; Rosenberg & Siegel, 2018), or compromise the experiment because it feels fun to do so (e.g., as for survey trolls; Lopez & Hillygus, 2018).

Many IBRs (called *quasi-controls*; Orne, 2009) were conducted in the sixties to deal with this general issue of *demand characteristics*, which can elicit *demand effects*. In these studies, participants received detailed information about an experimental procedure and were asked to complete measures or simulate behaviors as if they were subjected to the actual experimental treatment. For instance, behavioral responses to hypnosis were interpreted at the time as a demand effect because participants could convincingly simulate hypnotic behavior based on procedural information received about hypnotic treatments.

Being aware of the hypothesis tested in a study and being able to produce behaviors consistent with it is not inherently problematic. Consider for instance the Stroop effect: it is conceivable that participants may anticipate that it is easier to name the color of a word appearing in blue when this word is “Blue” than “Red”. This hypothesis-

awareness, however, would not rule out that the Stroop effect is produced by the operation of conflicting processes, nor would it imply that this effect is an experimental demand artifact.

Successful IBRs, however, raise the possibility that the effects of interest *may* be caused by demand characteristics. This likely arises when cues in the situation that are uncontrolled by the experimenter influence participants' interpretation of what responses are expected from them in the testing situation, as well as their motivation to produce responses consistent with this interpretation.

Consider the case of evaluative conditioning research. To remind the reader, an evaluative conditioning effect is found when the evaluation of a conditioned stimulus changes after its pairing with an affective stimulus (e.g., a formerly neutral kanji will elicit more positive evaluations after being paired with an affectively positive than negative stimulus). From the perspective of a participant taking part in an evaluative conditioning procedure (be it experienced or instructed), it should not be too difficult to figure out that evaluations (and which ones) are expected. Further, participants may be tempted to form impressions consistent with these expectations because producing them appears to be appropriate in the experimental setting.

For instance, in a study by Kurdi and Banaji (2017), "*participants were informed that they would see pairings of names with pictures such that one target group (e.g., Laapians) would always be paired with pictures of pleasant things and the other target group (e.g., Niffians) would always be paired with pictures of unpleasant things (...)*" (Kurdi & Banaji, 2017, p. 198). When receiving such instructions, participants may easily figure out that the experiment requires from them to start liking Laapians more than Niffians. When they form evaluations of the stimuli *in order to* comply/react with this perceived requirement of the testing situation, a demand effect is created.

Because compliance is key in characterizing demand effects, it is important to understand how it can operate and how it may be controlled for. In a recent theoretical review on demand characteristics (Corneille & Lush, 2023), the authors identified three compliance strategies: faking (intentional and conscious, involving no genuine experience), imagination (intentional and conscious, involving a genuine experience), and phenomenological control (akin to hypnotic suggestion; intentional and unconscious, involving a genuine experience; (Dienes, Lush, & Palfi, 2022; Dienes, Lush, Palfi, et al., 2022; Dienes & Lush, 2023). From that perspective, *hypothesis-aware* participants may produce either counterfeit or genuine responses because they want to produce them in the testing situation to meet the perceived requirements of the situation – sometimes, without being even aware that this is their intention (Dienes, Lush, & Palfi, 2022).

The case of phenomenological control is probably the most intriguing. To illustrate, participants taking part in a rubber hand illusion study may infer that the synchronous brushing of the fake hand with their real hand (concealed from their view) will produce the illusion of the rubber hand being their own. Once aware of this experimental hypothe-

sis, they may further want to produce responses consistent with it (i.e., to produce the illusion of the rubber hand being their own) because they think this experience is called-for in the testing circumstances and they want to comply with this understanding. However, there is a twist: because experiencing the illusion implies the phenomenological experience of producing it involuntarily, the illusion can be established only if participants inhibit their subjective experience of producing this illusion voluntarily. When participants can achieve that, the illusion is intentionally produced to comply with the requirements of the situation but without the accompanying experience of intentionality.

Phenomenological control, as a trait, is defined as the capacity to strategically experience as non-intentional perceptions, actions, or beliefs to meet the requirements of the situation. Individuals vary in their ability to play these meta-cognitive tricks on themselves, and this ability can be assessed with questionnaires of hypnotic suggestibility or derived from them, such as the Phenomenological Control Scale (Lush, Scott, et al., 2021) or the Sussex-Waterloo Scale of Hypnotisability (Lush et al., 2018). To sum up, when participants can implement phenomenological control over their subjective experiences, they can comply with the experimental hypothesis by producing responses consistent with it without being aware that this was their intention. This particularly tricky form of compliance generates genuine experiences without awareness of intentionality (for a more extensive treatment of how phenomenological control relates to compliance and experimental demands, see Corneille & Lush, 2023).

Controlling compliance with the experimental hypothesis, in its various forms, is critical to sound research. Such compliance is a major threat not just to the external validity of the study (i.e., will the effect generalize to other situations?) but also to its internal validity (i.e., is the interpretation of the effect correct?) For instance, participants may evaluate meaningless stimuli more positively when asked to approach rather than avoid them because they think this is the soundest behavior called for by the testing condition. If so, the effect of approach and avoidance on liking may not survive if changing cues in the situation that decrease participants' compliance motivation or their interpretation of the behavior called for by the testing situation (external validity issue). If the effect reflects compliance with the experimental hypothesis, this additionally suggests that it may be driven by mechanisms that differ from those assumed. As long as there are uncontrolled cues in the experimental situation that shape goals in the participants to produce the effect (or to be reactant to it), internal validity issues arise that contaminate the interpretation of the findings.

Different strategies may be considered for assessing or preventing compliance, but here too, they come with important caveats. Researchers may directly ask participants if they complied. However, this approach would only identify strategies based on imagination. Phenomenological control is supposed to be unconsciously implemented, and it may not be reported as a result. Turning to "faking" participants, by definition, one may not trust their responses.

Table 1. Questions raised by successful instruction-based replications

<i>What does psychological science tell us about effects beyond common knowledge?</i>
• How do individuals acquire psychological knowledge relating procedures to responses? • How does this knowledge differ from established psychological theories? • Do individuals rely on accurate mechanistic knowledge when forming psychological predictions? • Would instruction-based replications extend to moderating effects? • Can psychological science draw on this naïve knowledge for its theory building?
<i>Does performing the experienced version of the task add to the effect, how much so, and why?</i>
• Does performing the task contribute to the original effects, and to what extent? • Do distinct cognitive processes underly this distinct contribution? • How can we best estimate the comparability of effects found in experience-based and instruction-based replication studies? • Which psychological phenomena have been or may be successfully replicated on a verbal basis? • Do IBR outcomes depend on the measures of interest (e.g., simulation, prediction for self or others), on how the constructs are measured (e.g., directly or indirectly), and on individual differences?
<i>Should the effect be considered an experimental demand artifact?</i>
• How can demand effects be identified, assessed, controlled for, or eliminated? • To what extent are demand effects compatible or incompatible with current theorizing? • Can we draw on demand effects to enrich psychological theorizing and practice?

Researchers may also use measures less amenable to strategic control. This may work to some extent for faking, keeping in mind that both indirect behavioral measures (e.g., Cummins & De Houwer, 2021; Fiedler & Bluemke, 2005) and physiological measures (Damaser et al., 1963) can be faked. Regarding imagination and phenomenological control, they produce genuine responses, including physiological ones (Hägni et al., 2008); therefore, less controllable measures may not help (Corneille & Hütter, 2020; Corneille & Lush, 2023).

In the case of phenomenological control, we have seen above that it can be measured with individual difference questionnaires. Responses to phenomenological control questionnaires have been recently shown to predict a large diversity of effects, including the rubber hand illusion and mirror touch synesthesia (Lush et al., 2020), visually evoked auditory response, and autonomous sensory meridian response (ASMR) (Lush et al., 2022). Interestingly, it does not predict the Müller-Lyer illusion (Lush et al., 2022). It is less clear, however, how the tendency or ability to engage in faking and imagination as compliance strategies may be validly measured with individual difference questionnaires.

Finally, one may vary incentives for compliant behavior and see how this influences responses (De Houwer, 2022). This is a sound strategy, provided the (dis)incentives (1) are significant enough for overturning more powerful goals, (2) do not paradoxically exacerbate compliance/reactance, and (3) do not disrupt effects for reasons alien to compliance. For instance, promising big money to participants for taking the task seriously may induce faking (if one wants to be nice to the experimenter in return for the money) or reactance (if one feels bribed); it may also elicit task-related inferences that would otherwise not be drawn; and it may increase attention to the study materials.

More research is needed on compliance with experimental demand, and we are pleased to see a renewal of interest in this “scarecrow” (Coles et al., 2022; Corneille & Lush, 2023; Doyen et al., 2012; Forster et al., 2022; Klein et al.,

2012; Lush, 2020; Lush et al., 2021; Roseboom & Lush, 2022; Sharpe & Whelton, 2016). The lively debate currently taking place in research on bodily illusions is illustrative of the theoretical and methodological advances that research around these questions can generate (Chancel et al., 2022; Corneille & Lush, 2023; Ehrsson et al., 2022; Lush et al., 2020; Lush, Seth, et al., 2021; Lush & Seth, 2022; Reader, 2022; Roseboom & Lush, 2022; Slater & Ehrsson, 2022). Psychological reactance should not be overlooked either. For instance, recent research on the truth effect indicates that a sizeable portion of the participants do not show the effect, and sometimes even show a *reversed* effect (i.e., lower judgments of truth for repeated statements (e.g., Henderson et al., 2021; Schnuerch et al., 2021). This may signal measurement error. But this may also indicate that some participants drew reverse inferences, or were reactant to the veridical inferences they drew.

4. Looking Forward

IBRs raise outstanding questions for contemporary psychological research (for a summary, see [Table 1](#)). In this article, we asked: (1) To what extent does psychological knowledge go beyond common knowledge, (2) Does experiencing a task add to the effect of verbal instructions; and if so, is it via distinct processes, and (3) Are researchers studying phenomena artifactually inflated (or weakened) by participants’ compliance (reactance) with the perceived experimental hypothesis? We believe these questions represent opportunities for theoretical, methodological, and empirical development in psychological science.

Beyond advancing psychological theories and methods, IBRs can also inform practice. Whenever an experience-based (e.g., clinical) treatment is replicated based on mere verbal information, this points to a possible beneficial role of verbal suggestion. For instance, research suggests that fears acquired through Pavlovian conditioning can be reversed through verbal instruction about contingencies

(Mertens et al., 2018; Mertens & De Houwer, 2016). Likewise, social impressions thought to reflect deeply rooted mental associations can be quickly changed based on mere verbal information (Cone et al., 2021; Cone & Calanchini, 2021; Kurdi et al., 2022; Van Dessel, Ye, et al., 2019).

IBR research highlights the powerful role of causal inferences and expectancies in how people interpret their environment, including the perception of their self. In doing so, it connects to a long tradition of research on placebo effects (De Houwer, 2018a; Kirsch, 2018), hypnosis (Lynn et al., 2019), therapeutic compliance (Kanter et al., 2002, 2004), instruction-based learning (Kang et al., 2022), and contemporary research on predictive processing (Chancel et al., 2022; Clark, 2016; Martin & Pacherie, 2019; Van Dessel, Hughes, et al., 2019). IBR research may help identify cases where nondeceptive placebo treatments are effective. In nondeceptive placebo studies, patients receive verbal information about placebo effects and they are explicitly told that they will undergo a placebo treatment (for a recent discussion, see Colloca & Howick, 2018). These treatments alleviate ethical concerns inherent in placebo treatments, and their effectiveness has been supported in studies relying on self-reports and at least one neurophysiological measure (Guevarra et al., 2020). Imaginative suggestion has also been used for a long time for clinical treatment (Pintar & Lynn, 2008). Comparison of effect sizes between instruction-based and experience-based treatments may help identify when the latter should be used rather than the former.

5. Concluding Remarks

Successful IBRs have been documented for a variety of effects and they raise outstanding questions. However, we currently have little systematic information about which experience-based effects they do and do not replicate, and whether replication outcomes depend on the specific in-

stantiation of the IBR (e.g., simulation, prediction, or completion of the original measures; prediction for self or others; direct or indirect – including physiological – measures). In addition, among documented replications, information is scarce about how effects compare between experience-based versus instruction-based procedures, and this comparison is challenging. How compliance effects can best be addressed is also a thorny issue. This calls for more IBR research, and it is our hope that the ideas discussed here help promote it.

Author Contributions

Olivier Corneille: Conceptualization, Writing – original draft, Project administration, Writing – review and editing
Jérémy Béna: Conceptualization, Writing – review and editing

Competing Interests

The authors have no competing interests to declare.

Acknowledgments

Preparation of this article was supported by an FRS-FNRS grant (T.0061.18) awarded to Olivier Corneille, and by an FSR Incoming Postdoctoral Fellowship [IPSY FSR22 MOVE] awarded to Jérémy Béna. We are grateful for the comments we received from Tal Moran, Peter Lush, Patrice Terrier, Christoph Stahl, Marine Rougier, Florence Stinghamber, and Christian Unkelbach on a previous version of this manuscript, and for extensive email exchanges with Jan De Houwer. We thank Adam Hahn for giving us the impetus to write this article.

Submitted: April 24, 2023 PDT, Accepted: June 17, 2023 PDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

References

- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3), 183–200. <https://doi.org/10.1037/h0024835>
- Béna, J., Mauclet, A., & Corneille, O. (2022). Does co-occurrence information influence evaluations beyond relational meaning? An investigation using self-reported and mouse-tracking measures of attitudinal ambivalence. *Journal of Experimental Psychology: General*, 152(4), 968–992. <https://doi.org/10.1037/xge0001308>
- Béna, J., Melnikoff, D. E., Mierop, A., & Corneille, O. (2022). Revisiting dissociation hypotheses with a structural fit approach: The case of the prepared reflex framework. *Journal of Experimental Social Psychology*, 100, 104297. <https://doi.org/10.1016/j.jesp.2022.104297>
- Buettner, D., Nelson, T., & Veenhoven, R. (2020). Ways to Greater Happiness: A Delphi Study. *Journal of Happiness Studies*, 21(8), 2789–2806. <https://doi.org/10.1007/s10902-019-00199-3>
- Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, 65(1), 5–17. <https://doi.org/10.1037/0022-3514.65.1.5>
- Chancel, M., Ehrsson, H. H., & Ma, W. J. (2022). Uncertainty-based inference of a common cause for body ownership. *ELife*, 11, e77221. <https://doi.org/10.7554/elife.77221>
- Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190217013.001.0001>
- Coles, N. A., March, D. S., Marmolejo-Ramos, F., Larsen, J. T., Arinze, N. C., Ndukaihe, I. L. G., Willis, M. L., Foroni, F., Reggev, N., Mokady, A., Forscher, P. S., Hunter, J. F., Kaminski, G., Yüvrük, E., Kapucu, A., Nagy, T., Hajdu, N., Tejada, J., Freitag, R. M. K., ... Liuzza, M. T. (2022). A multi-lab test of the facial feedback hypothesis by the Many Smiles Collaboration. *Nature Human Behaviour*, 6(12), 1731–1742. <https://doi.org/10.1038/s41562-022-01458-9>
- Colloca, L., & Howick, J. (2018). Placebos Without Deception: Outcomes, Mechanisms, and Ethics. *International Review of Neurobiology*, 138, 219–240. <https://doi.org/10.1016/bs.irn.2018.01.005>
- Cone, J., & Calanchini, J. (2021). A Process Dissociation Model of Implicit Rapid Revision in Response to Diagnostic Revelations. *Personality and Social Psychology Bulletin*, 47(2), 201–215. <https://doi.org/10.1177/0146167220919208>
- Cone, J., Flaherty, K., & Ferguson, M. J. (2021). The Long-Term Effects of New Evidence on Implicit Impressions of Other People. *Psychological Science*, 32(2), 173–188. <https://doi.org/10.1177/0956797620963559>
- Corneille, O., & Hütter, M. (2020). Implicit? What Do You Mean? A Comprehensive Review of the Delusive Implicitness Construct in Attitude Research. *Personality and Social Psychology Review*, 24(3), 212–232. <https://doi.org/10.1177/1088868320911325>
- Corneille, O., & Lush, P. (2023). Sixty Years After Orne's *American Psychologist* Article: A Conceptual Framework for Subjective Experiences Elicited by Demand Characteristics. *Personality and Social Psychology Review*, 27(1), 83–101. <https://doi.org/10.1177/10888683221104368>
- Corneille, O., Mierop, A., Stahl, C., & Hütter, M. (2019). Evidence suggestive of uncontrollable attitude acquisition replicates in an instructions-based evaluative conditioning paradigm: Implications for associative attitude acquisition. *Journal of Experimental Social Psychology*, 85, 103841. <https://doi.org/10.1016/j.jesp.2019.103841>
- Cummins, J., & De Houwer, J. (2021). The Shape of Belief: Developing a Mousetracking-Based Relational Implicit Measure. *Social Psychological and Personality Science*, 12(8), 1517–1526. <https://doi.org/10.1177/1948550620978019>
- Damaser, E. C., Shor, R. E., & Orne, M. T. (1963). Physiological Effects During Hypnotically Requested Emotions. *Psychosomatic Medicine*, 25(4), 334–343. <https://doi.org/10.1097/00006842-196307000-00002>
- De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation*, 37(2), 176–187. <https://doi.org/10.1016/j.lmot.2005.12.002>
- De Houwer, J. (2018a). A Functional-Cognitive Perspective on the Relation Between Conditioning and Placebo Research. *International Review of Neurobiology*, 138, 95–111. <https://doi.org/10.1016/bs.irn.2018.01.007>
- De Houwer, J. (2018b). Propositional Models of Evaluative Conditioning. *Social Psychological Bulletin*, 13(3), e28046. <https://doi.org/10.5964/spb.v13i3.28046>
- De Houwer, J. (2022). On the Merits and Challenges of Treating Conscious and Unconscious Thoughts and Feelings as Behavior. *PsychArchives*. <https://doi.org/10.23668/PSYCHARCHIVES.5332>
- De Houwer, J., & Hughes, S. (2016). Evaluative Conditioning as a Symbolic Phenomenon: On the Relation between Evaluative Conditioning, Evaluative Conditioning via Instructions, and Persuasion. *Social Cognition*, 34(5), 480–494. <https://doi.org/10.1521/soc.2016.34.5.480>
- Dienes, Z., & Lush, P. (2023). The Role of Phenomenological Control in Experience. *Current Directions in Psychological Science*, 32(2), 145–151. <https://doi.org/10.1177/09637214221150521>

- Dienes, Z., Lush, P., & Palfi, B. (2022). Controlling Phenomenology by Being Unaware of Intentions. In J. Weisberg (Ed.), *Qualitative Consciousness* (1st ed., pp. 229–242). Cambridge University Press. <https://doi.org/10.1017/9781108768085.017>
- Dienes, Z., Lush, P., Palfi, B., Roseboom, W., Scott, R., Parris, B., Seth, A., & Lovell, M. (2022). Phenomenological control as cold control. *Psychology of Consciousness: Theory, Research, and Practice*, 9(2), 101–116. <https://doi.org/10.1037/cns0000230>
- Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral Priming: It's All in the Mind, but Whose Mind? *PLoS ONE*, 7(1), e29081. <https://doi.org/10.1371/journal.pone.0029081>
- Ehrsson, H. H. (2020). Multisensory processes in body ownership. In *Multisensory Perception* (pp. 179–200). Elsevier. <https://doi.org/10.1016/b978-0-12-812492-5.00008-5>
- Ehrsson, H. H., Fotopoulou, A., Radziun, D., Longo, M. R., & Tsakiris, M. (2022). No specific relationship between hypnotic suggestibility and the rubber hand illusion. *Nature Communications*, 13(1), 564. <https://doi.org/10.1038/s41467-022-28177-z>
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58(2), 203–210. <https://doi.org/10.1037/h0041593>
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and Unaided Response Control on the Implicit Association Tests. *Basic and Applied Social Psychology*, 27(4), 307–316. https://doi.org/10.1207/s15324834bas2704_3
- Fiedler, K., McCaughey, L., & Prager, J. (2021). Quo Vadis, Methodology? The Key Role of Manipulation Checks for Validity Control and Quality of Science. *Perspectives on Psychological Science*, 16(4), 816–826. <https://doi.org/10.1177/1745691620970602>
- Forster, P.-P., Karimpur, H., & Fiehler, K. (2022). Demand characteristics challenge effects in embodiment and presence. *Scientific Reports*, 12(1), 14084. <https://doi.org/10.1038/s41598-022-18160-5>
- Gast, A., & De Houwer, J. (2012). Evaluative Conditioning without Directly Experienced Pairings of the Conditioned and the Unconditioned Stimuli. *Quarterly Journal of Experimental Psychology*, 65(9), 1657–1674. <https://doi.org/10.1080/17470218.2012.665061>
- Gawronski, B., Brannon, S. M., & Bodenhausen, G. V. (2017). The associative-propositional duality in the representation, formation, and expression of attitudes. In R. Deutsch, B. Gawronski, & W. Hofmann (Eds.), *Reflective and impulsive determinants of human behavior* (pp. 103–118). Routledge/Taylor&Francis Group.
- Golaszewski, S., Frey, V., Thomschewski, A., Sebastianelli, L., Versace, V., Saltuari, L., Trinka, E., & Nardone, R. (2021). Neural mechanisms underlying the Rubber Hand Illusion: A systematic review of related neurophysiological studies. *Brain and Behavior*, 11(8). <https://doi.org/10.1002/brb3.2124>
- Guevarra, D. A., Moser, J. S., Wager, T. D., & Kross, E. (2020). Placebos without deception reduce self-report and neural measures of emotional distress. *Nature Communications*, 11(1), 3785. <https://doi.org/10.1038/s41467-020-17654-y>
- Hägni, K., Eng, K., Hepp-Reymond, M.-C., Holper, L., Keisker, B., Siekierka, E., & Kiper, D. C. (2008). Observing Virtual Arms that You Imagine Are Yours Increases the Galvanic Skin Response to an Unexpected Threat. *PLoS ONE*, 3(8), e3082. <https://doi.org/10.1371/journal.pone.0003082>
- Henderson, E. L., Simons, D. J., & Barr, D. J. (2021). The Trajectory of Truth: A Longitudinal Study of the Illusory Truth Effect. *Journal of Cognition*, 4(1), 29. <https://doi.org/10.5334/joc.161>
- Houston, J. P. (1983). Psychology: A Closed System of Self-Evident Information? *Psychological Reports*, 52(1), 203–208. <https://doi.org/10.2466/pr0.1983.52.1.203>
- Houston, J. P. (1985). Untutored Lay Knowledge of the Principles of Psychology: Do We Know Anything They Don't? *Psychological Reports*, 57(2), 567–570. <https://doi.org/10.2466/pr0.1985.57.2.567>
- Hu, X., Gawronski, B., & Balas, R. (2017). Propositional Versus Dual-Process Accounts of Evaluative Conditioning: II. The Effectiveness of Counter-Conditioning and Counter-Instructions in Changing Implicit and Explicit Evaluations. *Social Psychological and Personality Science*, 8(8), 858–866. <https://doi.org/10.1177/1948550617691094>
- Hütter, M., & De Houwer, J. (2017). Examining the contributions of memory-dependent and memory-independent components to evaluative conditioning via instructions. *Journal of Experimental Social Psychology*, 71, 49–58. <https://doi.org/10.1016/j.jesp.2017.02.007>
- Hütter, M., & Genschow, O. (2020). What is learned in approach-avoidance tasks? On the scope and generalizability of approach-avoidance effects. *Journal of Experimental Psychology: General*, 149(8), 1460–1476. <https://doi.org/10.1037/xge0000728>
- Kang, W., Pineda Hernández, S., Wang, J., & Malvaso, A. (2022). Instruction-based learning: A review. *Neuropsychologia*, 166, 108142. <https://doi.org/10.1016/j.neuropsychologia.2022.108142>
- Kanter, J. W., Kohlenberg, R. J., & Loftus, E. F. (2002). Demand characteristics, treatment rationales, and cognitive therapy for depression. *Prevention & Treatment*, 5(1). <https://doi.org/10.1037/1522-3736.5.1.541c>
- Kanter, J. W., Kohlenberg, R. J., & Loftus, E. F. (2004). Experimental and Psychotherapeutic Demand Characteristics and the Cognitive Therapy Rationale: An Analogue Study. *Cognitive Therapy and Research*, 28(2), 229–239. <https://doi.org/10.1023/b:cotr.0000021542.40547.15>
- Kasran, S., Hughes, S., & De Houwer, J. (2022). Learning via instructions about observations: Exploring similarities and differences with learning via actual observations. *Royal Society Open Science*, 9(3), 220059. <https://doi.org/10.1098/rsos.220059>

- Kelley, H. H. (1992). Common-Sense Psychology and Scientific Psychology. *Annual Review of Psychology*, 43(1), 1–24. <https://doi.org/10.1146/annurev.ps.43.020192.000245>
- Kilteni, K., Maselli, A., Kording, K. P., & Slater, M. (2015). Over my fake body: Body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in Human Neuroscience*, 9. <https://doi.org/10.3389/fnhum.2015.00141>
- Kirsch, I. (2018). Response Expectancy and the Placebo Effect. *International Review of Neurobiology*, 138, 81–93. <https://doi.org/10.1016/bs.irm.2018.01.003>
- Klein, O., Doyen, S., Leys, C., Magalhães de Saldanha da Gama, P. A., Miller, S., Questienne, L., & Cleeremans, A. (2012). Low Hopes, High Expectations : Expectancy Effects and the Replicability of Behavioral Experiments. *Perspectives on Psychological Science*, 7(6), 572–584. <https://doi.org/10.1177/1745691612463704>
- Kurdi, B., & Banaji, M. R. (2017). Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes? *Journal of Experimental Psychology: General*, 146(2), 194–213. <https://doi.org/10.1037/xge0000239>
- Kurdi, B., & Banaji, M. R. (2019). Attitude change via repeated evaluative pairings versus evaluative statements: Shared and unique features. *Journal of Personality and Social Psychology*, 116(5), 681–703. <https://doi.org/10.1037/pspa0000151>
- Kurdi, B., Mann, T. C., & Ferguson, M. J. (2022). Persuading the Implicit Mind: Changing Negative Implicit Evaluations With an 8-Minute Podcast. *Social Psychological and Personality Science*, 13(3), 688–697. <https://doi.org/10.1177/19485506211037140>
- Leynes, P. A., & Addante, R. J. (2016). Neurophysiological evidence that perceptions of fluency produce mere exposure effects. *Cognitive, Affective, & Behavioral Neuroscience*, 16(4), 754–767. <https://doi.org/10.3758/s13415-016-0428-1>
- Lopez, J., & Hillygus, D. S. (2018). Why So Serious?: Survey Trolls and Misinformation. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3131087>
- Lush, P. (2020). Demand Characteristics Confound the Rubber Hand Illusion. *Collabra: Psychology*, 6(1), 22. <https://doi.org/10.1525/collabra.325>
- Lush, P., Botan, V., Scott, R. B., Seth, A. K., Ward, J., & Dienes, Z. (2020). Trait phenomenological control predicts experience of mirror synaesthesia and the rubber hand illusion. *Nature Communications*, 11(1), 4853. <https://doi.org/10.1038/s41467-020-18591-6>
- Lush, P., Moga, G., McLatchie, N., & Dienes, Z. (2018). The Sussex-Waterloo Scale of Hypnotizability (SWASH): Measuring capacity for altering conscious experience. *Neuroscience of Consciousness*, 2018(1). <https://doi.org/10.1093/nc/niy006>
- Lush, P., Scott, R. B., Seth, A. K., & Dienes, Z. (2021). The Phenomenological Control Scale: Measuring the Capacity for Creating Illusory Nonvolition, Hallucination and Delusion. *Collabra: Psychology*, 7(1), 29542. <https://doi.org/10.1525/collabra.29542>
- Lush, P., Seth, A., Dienes, Z., & Scott, R. B. (2022). Trait phenomenological control in top-down and bottom-up effects: ASMR, Visually Evoked Auditory Response and the Müller-Lyer illusion. *PsyArXiv*. Preprint. <https://doi.org/10.31234/osf.io/hw4y9>
- Lush, P., & Seth, A. K. (2022). Reply to: No specific relationship between hypnotic suggestibility and the rubber hand illusion. *Nature Communications*, 13(1), 563. <https://doi.org/10.1038/s41467-022-28178-y>
- Lush, P., Seth, A. K., & Dienes, Z. (2021). Hypothesis awareness confounds asynchronous control conditions in indirect measures of the rubber hand illusion. *Royal Society Open Science*, 8(11), 210911. <https://doi.org/10.1098/rsos.210911>
- Lynn, S. J., Green, J. P., Polizzi, C. P., Ellenberg, S., Gautam, A., & Aksen, D. (2019). Hypnosis, Hypnotic Phenomena, and Hypnotic Responsiveness: Clinical and Research Foundations—A 40-Year Perspective. *International Journal of Clinical and Experimental Hypnosis*, 67(4), 475–511. <https://doi.org/10.1080/00207144.2019.1649541>
- Martin, J.-R., & Pacherie, E. (2019). Alterations of agency in hypnosis: A new predictive coding model. *Psychological Review*, 126(1), 133–152. <https://doi.org/10.1037/rev0000134>
- Mattavelli, S., Corneille, O., & Unkelbach, C. (2022). Truth by repetition ... Without repetition: Testing the effect of instructed repetition on truth judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <https://doi.org/10.1037/xlm0001170>
- Mattavelli, S., Van Dessel, P., & De Houwer, J. (2021). Why Does the IAT Reveal a Preference for Stimuli Said to Be Paired With an Unpleasant Sound? Stalking the Unexpected. *Collabra: Psychology*, 7(1), 18733. <https://doi.org/10.1525/collabra.18733>
- Merikle, P. M., & Reingold, E. M. (1991). Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(2), 224–233. <https://doi.org/10.1037/0278-7393.17.2.224>
- Mertens, G., Boddez, Y., Sevenster, D., Engelhard, I. M., & De Houwer, J. (2018). A review on the effects of verbal instructions in human fear conditioning: Empirical findings, theoretical considerations, and future directions. *Biological Psychology*, 137, 49–64. <https://doi.org/10.1016/j.biopsycho.2018.07.002>
- Mertens, G., & De Houwer, J. (2016). Potentiation of the startle reflex is in line with contingency reversal instructions rather than the conditioning history. *Biological Psychology*, 113, 91–99. <https://doi.org/10.1016/j.biopsycho.2015.11.014>
- Mertens, G., van Schie, K., Lammertink, S., Littel, M., & Engelhard, I. M. (2021). Verbal suggestions fail to modulate expectations about the effectiveness of a laboratory model of EMDR therapy: Results of two preregistered studies. *Journal of Behavior Therapy and Experimental Psychiatry*, 73, 101673. <https://doi.org/10.1016/j.jbtep.2021.101673>

- Moran, T., Van Dessel, P., Smith, C. T., & De Houwer, J. (2021). Can (Instructions About) Stimulus Pairings Influence Automatic and Self-Reported Evaluations in the Presence of More Diagnostic Evaluative Information? *Personality and Social Psychology Bulletin*, 47(8), 1249–1263. <https://doi.org/10.1177/0146167220964638>
- Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLOS Biology*, 18(3), e3000691. <https://doi.org/10.1371/journal.pbio.3000691>
- Nuel, I., Fayant, M.-P., Morgado, N., Subra, B., & Alexopoulos, T. (2022). The (Virtual) Reality of Social Approach-Avoidance Behaviours: Operationalisation Development and Construct Validity Testing. *Collabra: Psychology*, 8(1), 34197. <https://doi.org/10.1525/collabra.34197>
- Olsson, A., & Phelps, E. A. (2007). Social learning of fear. *Nature Neuroscience*, 10(9), 1095–1102. <https://doi.org/10.1038/nn1968>
- Orne, M. T. (2009). 5 Demand Characteristics and the Concept of Quasi-Controls. In R. Rosenthal, R. L. Rosnow, & A. E. Kazdin (Eds.), *Artifacts in Behavioral Research* (1st ed., pp. 110–137). Oxford University Press New York. <https://doi.org/10.1093/acprof:oso/9780195385540.003.0005>
- Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, 94(1), 16–31. <https://doi.org/10.1037/0022-3514.94.1.16>
- Pintar, J., & Lynn, S. J. (2008). *Hypnosis*. Wiley-Blackwell. <https://doi.org/10.1002/9781444305296>
- Reader, A. T. (2022). What Do Participants Expect to Experience in the Rubber Hand Illusion? A Conceptual Replication of Lush (2020). *Collabra: Psychology*, 8(1), 35743. <https://doi.org/10.1525/collabra.35743>
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7(1), 1–75. [https://doi.org/10.1016/1041-6080\(95\)90031-4](https://doi.org/10.1016/1041-6080(95)90031-4)
- Roseboom, W., & Lush, P. (2022). Serious Problems With Interpreting Rubber Hand “Illusion” Experiments. *Collabra: Psychology*, 8(1), 32274. <https://doi.org/10.1525/collabra.32274>
- Rosenberg, B. D., & Siegel, J. T. (2018). A 50-year review of psychological reactance theory: Do not read this article. *Motivation Science*, 4(4), 281–300. <https://doi.org/10.1037/mot0000091>
- Rougier, M., Muller, D., Ric, F., Alexopoulos, T., Batailler, C., Smeding, A., & Aubé, B. (2018). A new look at sensorimotor aspects in approach/avoidance tendencies: The role of visual whole-body movement information. *Journal of Experimental Social Psychology*, 76, 42–53. <https://doi.org/10.1016/j.jesp.2017.12.004>
- Rougier, M., Schmitz, M., & Yzerbyt, V. (2021). When my actions shape your looks: Experience-based properties of approach/avoidance bias the visual representation of others. *Journal of Personality and Social Psychology*, 120(5), 1146–1174. <https://doi.org/10.1037/pspa0000268>
- Schacter, D. L., Bowers, J., & Booker, J. (1989). Intention, awareness, and implicit memory: The retrieval intentionality criterion. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 47–65). Lawrence Erlbaum Associates, Inc.
- Schnuerch, M., Nadarevic, L., & Rouder, J. N. (2021). The truth revisited: Bayesian analysis of individual differences in the truth effect. *Psychonomic Bulletin & Review*, 28(3), 750–765. <https://doi.org/10.3758/s13423-020-01814-8>
- Sharpe, D., & Whelton, W. J. (2016). Frightened by an Old Scarecrow: The Remarkable Resilience of Demand Characteristics. *Review of General Psychology*, 20(4), 349–368. <https://doi.org/10.1037/gpr0000087>
- Shimizu, Y., Lee, H., & Uleman, J. S. (2017). Culture as automatic processes for making meaning: Spontaneous trait inferences. *Journal of Experimental Social Psychology*, 69, 79–85. <https://doi.org/10.1016/j.jesp.2016.08.003>
- Slater, M., & Ehrsson, H. H. (2022). Multisensory Integration Dominates Hypnotisability and Expectations in the Rubber Hand Illusion. *Frontiers in Human Neuroscience*, 16, 834492. <https://doi.org/10.3389/fnhum.2022.834492>
- Smith, C. T., Calanchini, J., Hughes, S., Van Dessel, P., & De Houwer, J. (2020). The impact of instruction- and experience-based evaluative learning on IAT performance: A Quad model perspective. *Cognition and Emotion*, 34(1), 21–41. <https://doi.org/10.1080/02699931.2019.1592118>
- Tsakiris, M. (2010). My body in the brain: A neurocognitive model of body-ownership. *Neuropsychologia*, 48(3), 703–712. <https://doi.org/10.1016/j.neuropsychologia.2009.09.034>
- Unkelbach, C., Koch, A., Silva, R. R., & Garcia-Marques, T. (2019). Truth by Repetition: Explanations and Implications. *Current Directions in Psychological Science*, 28(3), 247–253. <https://doi.org/10.1177/0963721419827854>
- Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition*, 18(1), 22–38. <https://doi.org/10.1016/j.concog.2008.09.006>
- Van Dessel, P., Cone, J., Gast, A., & De Houwer, J. (2020). The impact of valenced verbal information on implicit and explicit evaluation: The role of information diagnosticity, primacy, and memory cueing. *Cognition and Emotion*, 34(1), 74–85. <https://doi.org/10.1080/02699931.2019.1594703>
- Van Dessel, P., De Houwer, J., Gast, A., Roets, A., & Smith, C. T. (2020). On the effectiveness of approach-avoidance instructions and training for changing evaluations of social groups. *Journal of Personality and Social Psychology*, 119(2), e1–e14. <https://doi.org/10.1037/pspa0000189>
- Van Dessel, P., De Houwer, J., Gast, A., Smith, C. T., & De Schryver, M. (2016). Instructing implicit processes: When instructions to approach or avoid influence implicit but not explicit evaluation. *Journal of Experimental Social Psychology*, 63, 1–9. <https://doi.org/10.1016/j.jesp.2015.11.002>

- Van Dessel, P., De Houwer, J., Gast, A., & Tucker Smith, C. (2015). Instruction-Based Approach-Avoidance Effects : Changing Stimulus Evaluation via the Mere Instruction to Approach or Avoid Stimuli. *Experimental Psychology*, 62(3), 161–169. <https://doi.org/10.1027/1618-3169/a000282>
- Van Dessel, P., De Houwer, J., & Smith, C. T. (2018). Relational information moderates approach-avoidance instruction effects on implicit evaluation. *Acta Psychologica*, 184, 137–143. <https://doi.org/10.1016/j.actpsy.2017.03.016>
- Van Dessel, P., Hughes, S., & De Houwer, J. (2019). How Do Actions Influence Attitudes? An Inferential Account of the Impact of Action Performance on Stimulus Evaluation. *Personality and Social Psychology Review*, 23(3), 267–284. <https://doi.org/10.1177/1088868318795730>
- Van Dessel, P., Mertens, G., Smith, C. T., & De Houwer, J. (2017). The Mere Exposure Instruction Effect : Mere Exposure Instructions Influence Liking. *Experimental Psychology*, 64(5), 299–314. <https://doi.org/10.1027/1618-3169/a000376>
- Van Dessel, P., Ye, Y., & De Houwer, J. (2019). Changing Deep-Rooted Implicit Evaluation in the Blink of an Eye: Negative Verbal Information Shifts Automatic Liking of Gandhi. *Social Psychological and Personality Science*, 10(2), 266–273. <https://doi.org/10.1177/1948550617752064>
- Wisniewski, D., Braem, S., González-García, C., De Houwer, J., & Brass, M. (2022). Effects of experiencing CS-US pairings on instructed fear reversal. *Neuroscience*. Preprint. <https://doi.org/10.1101/2022.04.05.487162>

Supplementary Materials

Peer Review History

Download: https://collabra.scholasticahq.com/article/82234-instruction-based-replication-studies-raise-challenging-questions-for-psychological-science/attachment/169254.docx?auth_token=vlyuNAhBjPNIFo0MKcaE
