Causal Inference in Generalizable Environments: Systematic Representative Design

Lynn C. Miller, Sonia Jawaid Shaikh, David C. Jeong, Liyuan Wang,

University of Southern California

Traci K. Gillig,

Washington State University

Carlos G. Godoy, and Paul R. Appleby

University of Southern California

Charisse L. Corsbie-Massay

Syracuse University

Stacy Marsella

University of Glasgow and Northeastern University

John L. Christensen

University of Connecticut

Stephen J. Read

University of Southern California

Author Note

Lynn C. Miller, Sonia Jawaid Shaikh, David C. Jeong, Liyuan Wang, University of Southern California, Traci K. Gillig, Washington State University, Carlos G. Godoy, and Paul R. Appleby, University of Southern California; Charisse L. Corsbie-Massay, Syracuse University; Stacy Marsella, University of Glasgow and Northeastern

University; John L. Christensen, University of Connecticut; Stephen J. Read, University of Southern California.

Research reported in this article was supported by the National Institute on Drug Abuse under R01DA031626 awarded to Stephen Read (PI), by the National Institute of Mental Health under R01MH082671, awarded to Lynn Miller (PI), and the National Institute of General Medical Sciences under R01GM10996 awarded to Stephen Read and Lynn Miller (PIs). The content is solely the responsibility of the authors and does not necessarily represent the official views of NIDA, NIMH, or NIGMS.

Correspondence concerning this article should be addressed to Lynn C. Miller, University of Southern California, Annenberg School for Communication and Journalism, Los Angeles, CA 90089-1694. E-mail: lmiller@usc.edu

Abstract

Causal inference and generalizability *both* matter. Historically, systematic designs emphasize causal inference, while representative designs focus on generalizability. Here, we suggest a transformative synthesis – *Systematic Representative Design (SRD)* – concurrently enhancing both causal inference and "built-in" generalizability by leveraging today's intelligent agent, virtual environments, and other technologies. In SRD, a "default control group" (DCG) can be created in a virtual environment by *representatively sampling* from real-world situations. Experimental groups can be built with systematic manipulations onto the DCG base. Applying *systematic design features* (e.g., random assignment to DCG versus experimental groups) in SRD affords valid causal inferences. After explicating the proposed SRD synthesis, we delineate how the approach concurrently advances generalizability and robustness, cause-effect inference and precision science, a computationally-enabled cumulative psychological science supporting both "bigger theory" and concrete implementations grappling with tough questions (e.g., what is context?) and affording rapidly-scalable interventions for real-world problems.

Keywords: experimental design, representative design, systematic design, Brunswik, Systematic Representative Design, generalizability, virtual environments, games, cause and effect, virtual reality

Causal Inference in Generalizable Environments: Systematic Representative Design

Science and everyday life cannot and should not be separated. Science, for me, gives a partial explanation for life. In so far as it goes, it is based on fact, experience and experiment.

Rosalind Franklin, Ph.D.

Across science, "causality is at the crux of metaphysical, epistemological, and methodological issues" (Illari, Russo, & Williamson, 2011, p. 20). Although modern causal analysis, involving "mathematizing causality", using causal diagrams and symbolic languages (e.g., *do*-calculus) (Pearl & Mackenzie, 2018, p. 7), offers the capacity for valid cause-effect claims across a range of methods (Pearl & Mackenzie, 2018), historically, classic experiments have most readily afforded such inferences. But, "to different degrees, all causal relationships are *context dependent*[1], so the generalization of experimental effects [across studies, with different populations of individuals, settings, situations, methods, and so forth, also referred to as external validity] is always at issue" (Shadish, Cook, & Campbell, 2002, p. 5, material in brackets added, words italicized for emphasis).[2]

---

[1] Context matters; but, rarely is it defined. Van Overwalle (1997) draws on Rescorla and Wagner's (1972, p. 88) definition from the animal learning literature in which context refers to the relatively constant "situational stimuli arising from the ...environment" (see Van Overwalle & Van Rooy, 1998 for further treatment of this issue). For example, some aspects of context may provide a more or less constant background across individuals (e.g., a bar has different affordances than a church or a school). But, even there, other aspects of the "context" can be changing (e.g., whether we are in a "pick-up" situation or not and if so what is our role (given goals, beliefs, etc.) and where in the potential "pick-up" scenario are we)

[2] Rarely is a single factor,"X", deterministic (i.e., always necessary *and* sufficient) in causing "Y". Instead, causal relationships tend to be probabilistic (Eells, 1991), such that "many factors are usually required for an effect to occur, but we rarely know all of them and how they relate to each other." (Shadish et al., 2002, p. 5). Thus, across experiments "a given causal relationship will occur under some conditions but not universally across time, space, human populations, or other kinds of treatments and outcomes that are more

As Brewer and Crano (2014, p. 19) note, generalizability or replicability, could be designed into an experimental operationalization by using *representative* sampling of the targets to which researchers wish to generalize. Representative sampling (e.g., of persons, stimuli, contexts, and their interactions) in the design phase of developing one's experimental and control conditions would simultaneously "build in" a new type of generalizability – *generalizability to everyday life* (GEL). Although historically most researchers have deemed it not feasible within experiments to optimize representativeness of the "*organism-in-situation"* (Cronbach, 1957) to which we wish to generalize, the goal of the current work is to suggest that it is advantageous and feasible using today's technologies to build-in the capacity for *both* valid causal inferences and GEL into our experiments.

The argument that we should build the capacity for causal inference within generalizable environments fits with other emerging experimental paradigm shifts in the biological sciences. Motivated by the replicability crisis and a related issue, the often poor transferability of findings from mice to humans, biologists are developing new experimental designs "in more natural habitats [with wild mice that] can deliver results dramatically different from those in traditional laboratories – with profound implications for biomedical science" (Beans, 2018, p. 3196, material in brackets added). Rethinking traditional trade-offs in biology between control (e.g., controlled diet, temperature) and

---

or less related to those studied." (Shadish et al., 2002, p. 5). Furthermore, modern causal analysis has broad implications, including for the transportability of experimental results to new populations (e.g., in observational studies), or external validity (Pearl & Bareinboim, 2014; Pearl & Mackenzie, 2018). But, the generalizability of interest in these analyses (i.e., external validity) does not insure generalizability to everyday life (GEL).

generalizability (e.g., better transferability when wild mice live under more natural organism-environmental conditions), some argue a paradigm shift in experimental design is imminent (Garner, Gaskill, Weber, Ahloy-Dallaire, & Pritchett-Corning, 2017). Like the biologists above, we argue for new experimental paradigms in more real-world representative contexts with more "organism-in-situation" interactions (Cronbach, 1957, p. 682) to which we wish to generalize.

Both the capacity for valid causal inference[3] *and* generalizability[4] matter: But, why do psychologists need to worry about building causal inference capacity into *generalizable* environments? For many decades, researchers have complained about the generalizability of college samples (Gergen, 1973; Sears, 1986). Recently, psychology may have reached its own tipping point: A meta-analysis made it shockingly clear that WEIRD (Western, Educated, Industrialized, Rich, and Democratic) samples are often outliers and their behavioral and response patterns are among the least representative across human samples, raising concerns about making broad claims about human nature using WEIRD samples in (Henrich, Heine, & Norenzayan, 2010). Unfortunately, participant sampling is only part of the generalizability problem.

---

[3] Historically, the capacity for causal inference within experiments is defined in terms of a series of procedures that eliminate alternative explanations for differences in the dependent variable between, for example, the experimental and control groups, other than as due to the prior manipulation of the independent variable (see below in the section on systematic design and the promise of causality; see also, Brewer & Crano, 2014).

[4] Generalizability to everyday life (GEL) is a new concept that is defined as the "built in" capacity to generalize the results of a study. We do so by first identifying-- using formative research -- what for the persons of interest (POI) are the situations and sequences of interest (SOI) leading up to the behaviors of interest (BOI). Using sampling theory we representatively sample from these SOI for BOI to which we wish to generalize, *implementing* these SOI and BOI in the default control group (DCG), for example in a digital game (see Miller et al,, 2019; Appendix). External validity has historically been used as a measure of generalizability in experiments: It is defined in terms of the capacity to generalize the cause-effect relationship found in one study to the effects found in another with a different sample (e.g., of participants; or stimuli, etc.). External validity, however, does not insure GEL (see below for further discussion).

Currently, generalizability across experimental operationalizations of *social contexts* can also be problematic (Ceci, Kahan, & Braman, 2010). It is not surprising that context similarity was associated with the success of exact replication efforts in a sample of 100 replication study attempts (Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016). But, even with similarity in location, time, or culture, emerging examples from the psychological literature suggest that moving from more traditional, less representative stimuli and contexts to more naturalistic and representative contexts can *fundamentally alter* patterns of findings: The underlying mediational processes implicated for the target behavior of interest can be dramatically different (Gendron, Mesquita, & Barrett, 2013; Levine, Blair, & Clare, 2013).

How did we end up in this situation? Historically, the presumption was that "tight" experimental control (e.g., often using denatured stimuli or atypical contexts in operationalizations that might separate a potential "X" causal variable from others) enhanced internal validity (capacity to make valid causal inferences from "X" to "Y"). More stripped-down stimuli might afford more "clean" systematic manipulations with fewer potential third variables (affording an alternative explanation)[5] (Aronson, Ellsworth, Carlsmith, & Gonzales, 1990). Many argued that design elements that increased internal validity (and operationalizations that reduced third variable explanations) should be prioritized over generalizability (Berkowitz & Donnerstein, 1982; Calder, Phillips, & Tybout, 1983; Campbell, 1957; Mook, 1983). However, by generalizability here, what was meant was external validity (Campbell, 1957; Campbell

---

[5] But, because they were stripped down, they might produce smaller effects (i.e., such stimuli might be less involving and impactful) (Aronson, Ellsworth, Carlsmith, & Gonzales, 1990).

& Stanley, 1966). This involved demonstrating that a cause-effect relationship found in one, typically experimental study, is found when the population of individuals, settings, methods, and so forth are changed, often in a non-systematic, piecemeal fashion, with similar narrow operationalizations. The logic went: demonstrate cause-effect first, then external validity, and if need be, moderators. But, there is a fundamental problem here: Demonstrating external validity does *not* mean that these cause-effect relationships necessarily generalize to everyday life.[6]

Situations that are not, by design, representative in the first place, may demonstrate external validity (e.g., cause-effect relationship in one experiment found in a second with a different population) but not generalizability to everyday life (GEL): *The extent of the problem (*i.e., that our findings may not have GEL) *is unknown*. When operationalizations are unrepresentative, laboratory effects and processes may be their own kind of "weird" (Ceci, Kahan, & Braman, 2010). Social interactions, including across diverse groups/targets, are a critical part of the social context. However, social psychologists do not typically have participants interact with a range of partners in experiments, with few exceptions (e.g., speed dating, see Finkel & Eastwick, 2008), let

---

[6] A manipulation's experimental realism and social impactfulness may enhance effects but may add extraneous variables: But, reducing extraneous variables (adding control) may weaken (potentially reducing replicability of) experimental effects (Aronson, Ellsworth, Carlsmith, Gonzales, 1990). Field research, with greater naturalism, is one compromise in this tension, but it neither necessarily affords the *representative* situational/setting and samples to which we would like to generalize nor the experimental control typically found in a laboratory study (Paluck & Cialdini, 2014). Furthermore, an event in the lab may be similar to an event in real life (exhibiting mundane realism) (Aronson et al., 1990) without it having GEL. To insure that the effects and relationships (e.g., among variables) in the control are generalizable to everyday life, the virtual environment (described later) by design needs to be representative of the cues and settings/situations of interest (SOI) likely to lead to the behavior of interest (BOI) for the population of interest (POI) as in everyday life. And, a condition with high experimental realism (e.g., some conformity studies (Aronson et al., 1990)) may include situations that aren't representative of challenging conformity situations, given the BOI, that individuals in a POI may confront in everyday life.

alone representative partners and interactions. Often a task (e.g., in social interaction; social perception) is heavily and unnaturally constrained: This may enhance experimental control but reduce GEL. For example, meta-analyses of experiments involving stimulus videos of students instructed to lie or tell the truth, show consistent evidence across studies that participants are not particularly good at detecting deception (Bond & DePaulo, 2006), but this may not generalize to other contexts, especially natural contexts where deception detection has been shown to be more frequent (Levine et al., 2013). In addition, psychologists' methodological toolbox does not include validity checks to assess how representative or weird our conditions (and operationalizations) might be compared to real-world phenomena, processes, and target behaviors of interest (BOI) for the target population of interest (POI). Furthermore, complacency about real-world generalizability undermines psychology's relevance to the public (Cialdini, 2009). Federal agencies (e.g., the Defense Advanced Research Projects Agency (DARPA), National Institutes of Health (NIH), National Science Foundation (NSF)) increasingly seek designs that afford *both* causal inference *and* real-world generalizability to serve public needs (Davis, O'Mahony, Gulden, Osaba, & Sieck, 2018).

Could an experimental paradigm, affording valid causal inferences, also optimize GEL? We aim here to address this question offering a new paradigm -- Systematic Representative Design (SRD) -- that is a synthesis of two major designs: representative design (Brunswik, 1943; Brunswik, 1955a, 1955b) and classic experimental or systematic designs (Shadish et al., 2002) with roots in Wundt (1902). Systematic designs prioritize designing in the capacity to make valid cause-effect inferences, while representative designs prioritize designing in the capacity for inferences about real-life generalizability. Historically, integrating these strengths seemed impractical. We argue that this is now feasible, partly due to the availability of enabling technologies. For example, technologies such as intelligent agents in narrative games allow static user and partner features (e.g., appearance) to be systematically altered, and other parameters (e.g., agent goals and beliefs; agent "theory of mind" complexity) can be differentially set within and across studies to systematically alter agent behaviors with great precision and replicability, creating an array of representative interaction partners and interactions. Thus, it is possible today to generate representative social interactions with diverse others in representative narratives (Miller et al., 2011; Miller et al., 2019).

In making our argument, we first review the key strengths and weaknesses of systematic and representative designs. Second, we explicate the proposed Systematic Representative Design (SRD) synthesis. For this synthesis we use classic systematic experimental design features (e.g., random assignment). At the same time, we build in GEL for our target populations via task analysis, sampling, leveraging new technologies, and using correlations between virtual and real-world behavior (i.e., virtual validity checks) to assess the achievement of GEL for our default control group (DCG). Then, we design experimental groups based on that DCG base for systematic experimental comparisons. Third, we discuss how SRD concurrently advances generalizability and robustness, cause-effect inference and precision science, and a computationally-enabled cumulative psychological science supporting both "bigger theory" and concrete implementations grappling with tough questions (e.g., what is context?) and affording rapidly scalable interventions for real-world problems.

Systematic and Representative Designs

Below, we describe and compare systematic and representative designs (Table 1). Each offers distinct promises and challenges concerning causality and generalizability and both present additional common challenges that we also address.

Systematic Design: The Promise of Causality and the Challenge of Generalizability

Classic experimental designs, sometimes referred to as systematic designs (Brunswik, 1947), afford considerable strengths, but also have – at least as typically operationalized – considerable weaknesses. Procedurally, they can afford valid *causal claims* (i.e., that "X" causes "Y"). It has been argued that experimental designs also can provide *explanatory* inferences, that is, how, why, and under what conditions "X" causes "Y" (Brewer & Crano, 2014).

Criteria for the capacity to make valid causal inferences. All of the following criteria must be met to achieve the capacity to infer valid cause-effect relationships in experiments: (a) the temporal order of variables is such that the potential cause or independent variable (IV) precedes the effect or dependent variable (DV), (b) cause and effect covary with one another, and (c) the elimination of plausible alternative explanations (e.g., participants are randomly assigned to conditions). These criteria largely ensure that the cause-effect relationship claims are not undermined (Brewer & Crano, 2014). Random assignment of participants is a powerful systematic design feature because every participant has the same chance of being assigned to any given condition (e.g., control versus experimental) and differences between conditions cannot be attributed to third variables related to the participants (e.g., participant propensities). In contrast, self-selection into a condition could result in a third (extraneous) variable responsible for outcome differences between conditions.[7] Within a given research setting, individual studies using these procedures are said to afford causal inferences with high internal validity (Campbell, 1957; Campbell & Stanley, 1966). While criteria (a) and (b) above do not *require* an experiment – a correlational study could suffice[8] – experiments readily meet all three requirements.

Operationalizations and construct validity. Despite these strengths for making cause-effect inferences, GEL of cause-effect inferences does not depend upon the three basic procedural criteria or elements of systematic designs mentioned above. Rather, it can be profoundly affected by how we operationalize our variables. Operationalizations from a single study may vary in the extent to which they adequately represent the theoretical constructs or processes of interest (i.e., construct validity) (Brewer & Crano, 2014). For example, findings in research on emotion cue judgment suggest that the naturalness of stimuli and contexts in the real-world (e.g., static faces in emotion cue judgment research versus faces in motion or faces with bodily cues) does not merely

---

[7] As some experiments, *within* a condition, "self-selection" occurs, especially in long interventions. An advantage of the proposed SRD is that these self-selection opportunities *within* conditions can be *kept constant across conditions* or systematically manipulated.

[8] In correlational designs, there can be a predictor or measured independent variable "X" that *precedes* a dependent variable "Y," where there is sufficient variability on both variables and where X and Y significantly covary. Such a measured independent variable typically involves naturally occurring behavioral variations that are then correlated with the dependent variable. This provides a necessary, albeit not sufficient basis, for causal inference (Brewer & Crano, 2014). Causal analysis (e.g., Pearl & Mackenzie, 2018) suggests, however, greater causal inference potential from correlational data, especially for "big data" over time.

moderate emotion cue judgment, it alters neural processing in those judgments. Indeed, context is always present, regardless of the researcher's intention or awareness, and is arguably "integral to the emotional perception itself" (Gendron et al., 2013, p. 6). Another example of how operationalizations can go astray involves deception detection research where, in the typical design, experimentalists manipulate instructions across conditions, then assess human capabilities to detect others' lies. "Senders" (e.g., undergraduate participants) are *asked* to lie or tell the truth, then "receivers" must try to discriminate lies from truths, judging only the verbal and non-verbal cues of the target on a videotape. A meta-analysis investigating 206 documents indicates, "people achieve an average of 54% correct lie-truth judgments, correctly classifying 47% of lies as deceptive and 61% of truths as nondeceptive" (Bond & DePaulo, 2006, p. 214). That is, this near chance hit rate meant that "many lies are undetectable" (Bond & DePaulo, 2006, p. 231). Despite cross-study consistency, Bond and DePaulo were concerned about the denatured context of typical lab settings, the information available to receivers (and the information that was withheld from them) in making lie determinations, and the impact of that on real-world generalizability. Partly to address this, Levine et al. (2013) created a more naturalistic laboratory situation in which initial participants (IPs) could decide to go along with (or not) a confederate who wished to cheat on a trivia game. Afterwards, an expert interviewer from the Federal Bureau of Investigations (FBI), blind to whether the initial participants (IPs) cheated or not, questioned and prompted diagnostic information from each of the IP's in a naturalistic context: The entire interaction was recorded and shown to new participant observers (POs), similar to the experience of a court trial. Levine et al.

(2013) found that new POs, blind to the IPs' cheating, who saw a random sample of 36 expert interviews had high detection deception accuracy (i.e., almost 94%). Thus, more naturalistic set-ups (e.g., faces naturally occurring with the body versus not; deception-detection where target deception opportunities produce cues affording detection versus not) can dramatically alter researcher's causal inferences**.**

Representative Design: The Promise of Generalizability and the Challenge of Causality

Before discussing representative design, we introduce its theoretical foundation, probabilistic functionalism (see Dhami, Hertwig, & Hoffrage, 2004 for a review). Then, we introduce representative design as the methodological companion for this theory.

Probabilistic functionalism. Brunswik's (1952) Darwinian functional approach, *probabilistic functionalism,* assumed that the environment for the organism is uncertain: To adapt, the organism must learn, not necessarily consciously, to achieve the organism's goals (distal object) using environmental (proximal) cues that provide only probabilistic indicators (Hammond & Stewart, 2001). Thus, psychological researchers should aim to discover probabilistic laws describing an organism's adaptation (distal achievement) to the "causal texture of its environment" (Dhami et al., 2004, p. 962). They could do that by asking, "how is an organism perceiving and responding to its probabilistic environment to achieve a distal variable? Can the findings of such an experiment be used to predict future achievement in that environment?" (Dhami et al., 2004, p. 962).

Visual depiction of probabilistic functionalism. Brunswik's (1952) theoretical model has been referred to as "the Lens Model" given its visualization (see Figure 1). To illustrate, we use an example from Dhami and Belton (2017) and start with a distal (or in Figure 1, environment) criterion (e.g., the public's perception that judges are being fair). Imagine an available set of proximal cues in the environment and their inter-cue correlations that might optimally predict perceptions of a judge's fairness. *Ecological validities*[9] refer to coefficients indicative of proximal environmental cue validity in predicting to the specified distal criterion (environment) state or policy (Araújo, Davids, & Passos, 2007). Taking into account a judge's inter-cue correlations, these proximal cues and their relative weights, could be gleaned from that judge's prior experience in making decisions in a probabilistic environment.[10] As Gigerenzer, Hoffrage, and Goldstein (2008) note, Brunswik referred to the cues in the judge's mind and how he or she uses those cues to make a judgment as the *cue utilization validities* (see Figure 1) and argued for exploring and potentially statistically controlling for covariates (i.e., inter-cue correlations), preferring a correlational/partial-correlation approach.[11] Furthermore, these authors also note that Brunswik specified an achievement index, or a coefficient for the relationship between the optimized prediction that one might get from available cues in the environment and the judgment provided based on which and how cues were utilized.

From Brunswik's probabilistic functionalism, as Dhami et al. (2004) note (see also Hammond & Stewart, 2001), various system designs emerge that are presumably

---

[9] Brunswik's (1956) *ecological validity (actually Brunswik refers to the validities of the cues)* and *representative design* are often used incorrectly and interchangeably (Araújo et al., 2007). Here, we use them as Brunswik intended.

[10] Trying to learn the ecological validities of these cues, and their intercorrelations, via experience was presumed, with the organism ideally learning the equivalence and substitutability of different cues (e.g., if one set of cues were unavailable or unreliable).

[11] Dhami and Belton (2017) argue that their Brunswikian idiographic assessments and their heuristics either match or outperform alternatives relevant to cue utilization in making decisions.

called systems in part because feedback (e.g., from learning, partners) was assumed. For example, “the single system” approach involves using proximal cues to predict to the distal stimulus/criterion (or policies) and is the most common, especially in the social judgment policies domain. The “double-system design” (full Lens Model parameters, see Figure 1) has also been used in the social judgment domain (Dhami et al., 2004). The “triple-system design” involves two-person use of shared probabilistic cues and is used for studying interpersonal processes, whereas a “four-system design” is used for studying group processes (see Dhami et al., 2004; Dhami & Olsson, 2008; Hammond, 1965). One way to study such naturally occurring cues and their role in probability judgements that guide the behavior of interest (BOI) (e.g., perception, decision-making, behavior) is to reproduce the psychological gist of the real-world adapted-for environment, bringing it into the lab. But how?

Representative design: A methodological companion. Representative design was Brunswik’s innovative methodology that fit with his *probabilistic functionalism* approach. For Brunswik (1955b; 1956), studying these “*organism-environment relations”* (Dhami et al., 2004, p. 959) meant representatively sampling from the environmental situations of interest (SOI) and implementing these in a laboratory setting in the study design phase of research. The SOI include the settings, contextual features, narratives, situations, stimuli, and choices that might lead to a BOI for a given POI and to which the researcher wishes to generalize. Researchers can only make claims about the extent to which a certain phenomenon (e.g., a probabilistic cause-effect relationship in the real-world) occurs for a given population if there is adequate random sampling of situational cues for the BOI (Brunswik, 1955a; Brunswik, 1955b).[12] Brunswik urged psychologists to specify in the design phase, “To what circumstances do we wish to generalize, or apply, our results?” (Araújo et al., 2007, p. 72).

To bring the gist of the real-world into the lab, Brunswik pioneered human task analysis, with participants in his studies estimating the size of an object (based on some predetermined parameters) at random time intervals over a four-week period while experimenters objectively measured the object, repeating this procedure over 180 situations (see Dhami et al., 2004). This procedure provided the data Brunswik needed for designing his laboratory experiments’ perceptual stimuli for designed in generalizability. Such adequate situational sampling frames were and are feasible (Gigerenzer, Hoffrage, & Kleinbölting, 1991).[13]

Comparing Systematic and Representative Designs (see Table 1)

Systematic design is focused on establishing causal connections and uncovering “laws” using these key “designed in” features (e.g., random assignment) to enhance internal validity. In contrast, Brunswik’s goal was to discover, using primarily correlational methods (with statistical controls in the analysis phase), individuals’ probabilistic laws in making and adapting predictions using environmental cues to achieve real-world goals (Dhami et al., 2004). That is why an important goal in representative designs is to feature variables (e.g., cues) that are chosen for their importance in the environment (representativeness) and naturally “tied.” For his

---

[12] Nonprobability sampling techniques might require a validity check since they might not cover the ecology to which the researcher wished to generalize (Brunswik, 1955b; Dhami et al., 2004).

[13] Gigerenzer et al. (1991) showed that representative sampling of cities (vs. not) could dramatically impact inferences about bias, consistent with claims that participants’ use probabilistic mental models in making these inferences and judgments.

perceptual studies, Brunswik (1944) took the naturally occurring covariations (or "ties") that existed between the organism and environment in the field and recreated these in the lab. Alternatively, systematic designs' variables that are artificially "tied" in the lab, are not tied in naturalistically occurring contexts, and conversely variables are "untied" in the lab that are tied naturally in the real-world. For instance, researchers may provide a face without the rest of a body as a stimulus, thereby "untying" naturally occurring body parts [14]. Similarly, a researcher may manipulate whether someone is asked to lie or tell the truth without regard to whether this instruction naturally occurs, and if so, whether it does so with the same frequency and whether it covaries with expressed deceptive behavior. Thus, variables are not necessarily relevant to the natural ecology in systematic designs making them artifactually "untied" (or "tied") and limiting researchers' ability to assess psychological processes as they function for organisms in the environment for which the organism is adapted, i.e., reducing generalizability (Dhami et al., 2004).

Systematic and representative designs have complementary strengths and weaknesses: The former more readily affords the capacity for valid experimental causal inferences and the latter more readily affords the capacity for generalizability. However, currently neither researchers taking a systematic design approach nor researchers taking a representative design approach seriously consider *context*. Yet, doing so, and trying to address the question of, "what is the context here," is essential to adequately implement SRD. Context is often defined in terms of situations (e.g., Van Overwalle & Van Rooy, 1998). Unfortunately, we still do not have adequate definitions and taxonomies of those

---

[14] We can not even assume that across studies with faces, the same cause-effect relationships will result. For example, the orientation of the face can affect social perception (Witkower & Tracy, 2019).

either (see below). Understanding how people think about situations is key to understanding how humans structure social meaning and therefore, key to our ability to create SRDs. Of course, it is not enough to understand how individuals categorize only the current situation, we need to understand when, why, and how for whom the situation changes. Narratives, as we argue below, help **s**tructure understanding and prediction of social interaction within a situation and across situations over time.

Missing Pieces: Structuring the Dynamics of Social Interaction

What is a situation? Social situations, including the presence or implied presence of others (Allport, 1968), and a taxonomy of them, have long been thought to be critical to the very definition of social psychology (Baron, Byrne, & Suls, 1989; Hilton, 2012; Milgram, 1965). For example, Milgram said, "Ultimately, social psychology would like to have a compelling *theory of situations* which will, first present a language in terms of which situations can be defined; proceed to a typology of situations; and then point to the manner in which definable properties of situations are transformed into psychological forces in the individual*"* (1965, p. 74). But, despite advances in defining situations (Argyle, Furnham, & Graham, 1981; Rauthmann et al., 2014; Yang, Read, & Miller, 2006), ***there is neither consensus on what a situation is nor a "matrix" or "taxonomy" within which these "situations" or combinations of features composing situations are arranged***[15], which would tell us something fundamental about how changes in key

[15] This situation may be analogous to the situation in chemistry before the development of The Periodic Table that is used as the basis for predictions, hypothesis testing, and theory in chemistry that some have argued moved chemistry from a pre-paradigmatic science to a science as physicists would think of it (see Scerri, 2007).

underlying parameters alter behavior (Kenny, Mohr, and Levesque, 2001; Reis, 2008; Swann & Seyle, 2005).

However, agreement regarding defining situations, or identifying situations' underpinning features, may be emerging. For example, key to situations are their affordances for goal pursuit (Argyle et al., 1981; Baron & Boudreau, 1987; Argyle, Furnham, & Graham, 1981; Grant & Dweck, 1999; Miller, Cody, & McLaughlin, 1994; Miller & Read, 1991; Read & Miller, 1989a,1989b; Reis, 2008; Yang et al., 2006).
One way that the term situation has been construed is in terms of settings and situational awareness in those settings (Killingsworth, S. A. Miller., & Alavosius, 2016). A setting, like a church or an emergency room, constrains the likely goals and tasks, as well as likely scenarios and actions. Endsley (1995) conceptualized situation awareness in terms of three levels of consideration: (a) perception, (b) comprehension, and (c) projection. At the perceptual level, there is a focus on perceptual cues and stimuli that are relevant to successfully performing tasks or understanding the situation (e.g., if changes are slow or rapid); Brunswik's theory and approach were most focused at this level (Brunswik, 1955b). At the comprehension (or meaning) level, humans learn the organizing structures (e.g., narratives, scripts, pattern matches) to make sense of what is happening (e.g., an emergency). At the projection level, humans anticipate and predict future situations (e.g., given a shooting and numerous causalities, a hospital low on staff) based on what is happening and the rate with which it is changing. Top-down models and organizing structures can facilitate this future projection.

What is happening here over time? Why and how do I respond to it? Social interactions are complex and dynamic. For example, to understand how people extract meaning from the current, as well as prior and potentially future interactions, and then to respond to the situation, draws us into diverse literatures (e.g., in contextual sequence analysis, see Cornwell, 2015). Participants in our studies, as well as intelligent agents in some games (Marsella, Pynadath, & Read, 2004), are trying to do two things concurrently: (a) understand "what is happening here" (e.g., social perception, comprehension-meaning; projection into the future using theory of mind about self and other and prior experience/learning), and (b) understand "why and how should I respond to it" (decision-making and enactment of behavior). This is based on often automatic inferences involving theory of mind (e.g., about beliefs about self, other (e.g., Marsella, Pynadath, & Read, 2004)) and considerations of one's own goals and plans, as well as resources to achieve them, given prior experiences and what the individual/agent thinks might happen if he/she made various choices now and with unfolding choices by all parties involved.

In probabilistic functionalism, Brunswik focused on how people use perceptual cues to predict what was going to happen, but perceptual cues are unlikely to suffice in complex social interaction without a way to structure such data. How could streams of cues be structured to constrain and make sense of these cue utilizations over time? Brunswik and his followers did not answer that question. Others, especially those concerned with narrative structure, however, have tried to address it.

Narratives help structure understanding and prediction of social interaction over time. Stories or narratives provide a coherent way to understand the meaning of an action by *contextualizing* or *situating* the action in relation to the other actions or events involving self and others (e.g., Pennington & Hastie, 1986). Whether involving human behavior (e.g., Barker, 1963; Barker & Wright, 1955; Forgas, 1979; G. A. Miller, Galanter, & Pribram, 1960) or text processing (e.g., Mandler, 1978; Rumelhart, 1977), most action sequences have four components, including the sequence's goal, actions making up the plan to achieve the goal, initiating conditions, and the outcome. Evidence suggests that this source-goal-plan-unit structure facilitates memory (Abbott & Black, 1986). As Read (1987) notes, the action sequence alone is insufficient to provide this information; additional detailed information is required (e.g., knowledge of the actors' goals and how these actions fit together into a plan for goal achievement) or inferred, which forms the basis of a coherent mental representation that can be used to answer questions about what happened and why (e.g., Black, Galambos, & Read, 1984; Graesser, 1981; Schank & Abelson, 1977). Read's (1987) narrative-based model of attribution argued that humans use social knowledge structures (e.g., scripts, plans, goals, and themes) in constructing causal scenarios for causal reasoning and explanation. Often these structures are hierarchically nested (Figure 2), such that subgoals are achieved in the service of the goals most heavily activated in a given moment in the interaction, due to both within-person changes (e.g., due to situational or interoceptive state changes) (Read & Miller, in press; Read, Smith, Droutman, & Miller, 2017) and between-person differences (e.g., in chronic relative goal activations) (Read & Miller, in press).

Communication patterns and narratives can help frame the causal structure of social interaction. In social interactions, humans or actors typically take "turns" (e.g., alternating speaker and receiver roles). These "turns" of one social actor vis à vis another could potentially signal new situations and pertinent cues that involve inferences about "why" (e.g., Why did person A do or say that to Person B? Why did person B grimace when A said that to him/her?). As suggested above, these inferences leverage knowledge structures, (e.g., self and others' goals and beliefs; Miller & Read, 1987, 1991): Narratives or stories can frame and structure causal scenarios and meaning (Read & Miller, 1995; Schank & Abelson, 1977; Schank & Abelson, 1995). Indeed, narratives, which can facilitate interpersonal as well as group communication, may be so fundamental because of their capacity to fulfill core social motives, which are tied to identity (Costabile, Shedlosky-Shoemaker, & Austin, 2018).

Identity and cultural framing of the meaning of actions. Developmental and personality psychologists have argued that cultural and self-understanding (Bruner, 1987; Bruner, 1990; Fivush, Habermas, Waters, & Zaman, 2011; McAdams, 1990, 2001, 2008; Nelson & Fivush, 2004; Thorne & McLean, 2003) is not just about sequences of actions, but importantly about how action sequences are framed (e.g., in terms of the goals, beliefs, meaning of what just happened). Parental scaffolding of children's narrative framing of the world starts early--by 16 months of age (Reese, 2002). The neural basis of these stories about self and others, often implicating values, is gaining attention (e.g., D'Argembeau et al., 2014). Narrative is critical to the construction of meaning and therefore critical to representative design in the study of human behavior: Narrative ties together one's immediate and long-term experience pertaining to self- and cultural-identities, and projects expectancies regarding self, others, and the unfolding situation into the future.

Systematic Representative Design (SRD): A Unifying Design Framework

Systematic Representative Design (SRD) is an experimental approach that attempts to optimize the strengths and mitigate the weaknesses of systematic and representative designs, better affording *both* valid causal inferences and greater GEL (see Table 1). In addition, SRD uses representative narrative structures and other considerations (e.g., cueing based on cultural scripts) to grapple with "missing pieces" in past experimental designs that do not adequately contextualize the POI's everyday experiences and behavioral choices. To achieve this end, we leverage the power of new enabling technologies including virtual environments and games, intelligent agents, mobile technologies and sensors. Below, we first discuss the features that SRD adopts from standard systematic and representative designs, as well as additional innovations in research design. Then, we discuss these enabling technologies in the following section and how they enable SRD.

Optimizing Systematic Design Strengths in SRD

A strength of systematic designs, compared to representative designs, is the ability to more easily afford valid cause-effect relationship inferences. SRDs leverage those strengths by meeting the standard criteria of classic systematic design experiments. For example, in a standard two-cell between-subjects experimental design, SRD participants would be *randomly assigned* to the "default control group" (DCG) or one alternative experimental condition built upon the DCG base (E_1DCG). The experimental and control groups would be the same, except for the manipulation of the independent variable (e.g., in a two-cell design) or independent variables (in a multifactor design) after which the dependent BOIs are assessed. Therefore, differences in participant behavior after random assignment between conditions can be attributed to the manipulation of the independent variable. Thus: (a) the potential cause *precedes* the effect, (b) cause and effect covary with one another, and (c) important plausible alternative explanations are eliminated by random assignment, helping ensure that the conditions do not differ from one another except on the variables manipulated. The difference between an SRD and a systematic design is the DCG, which is designed to be representative of the settings, situations, and stimuli to which we wish to generalize for the POI and target BOI. [16]

Optimizing Representative Design Strengths in SRD

In systematic designs, neither control nor experimental groups are typically designed to be representative of the SOI to which the researcher wishes to generalize for the POI. In representative designs, researchers typically attempt to representatively

---

[16] Procedurally, in classic experimental designs, participants do not self-select themselves into condition. Instead, they are randomly assigned to conditions, to eliminate the possibility that differences between conditions are due to participants' pre-existing differences instead of the manipulation. Procedurally, in SRD, participants are randomly assigned to the DCG or a given experimental condition; participants then make choices (e.g., in a virtual environment game) just like in classic designs. Others (e.g., virtual intelligent agents) subsequently respond to the participants' responses, and those responses are adjusted given the participant's behavior. Different agent responses can then affect the participant's subsequent options and choices, creating interactional sequences into which participants "self-select" that are narratively designed to be more generalizable to everyday life (see below; Appendix; Miller, Wang, Jeong, & Gillig, 2019). However, there are not *between group differences* in self-selection opportunities in SRD, unless systematically manipulated by design with all other variables held constant or controlled as a source of alternative explanations. Thus, this type of "self-selection" during the course of a game does not present a threat to internal validity.

sample SOI and then assess cue utilizations in predicting some decision or response outcome (e.g., Gigerenzer et al., 1991). Research using representative designs has mostly been conducted in domains involving perception, decision-making, or social perception and judgment. In such contexts, there is typically one decision (e.g., setting bail) for a given individual (e.g., a judge in a given case). Task analysis enabling representative designs and conceptualizing cue utilization with new heuristics can require considerable formative work, but the payoff may be more evidence-based psychological models in line with how humans make decisions and respond to their actual environment. Such models can be more predictive and useful compared to alternative heuristics and statistical models (Dhami & Belton, 2017). However, past efforts to create representative designs typically lack extended dynamic social interaction leading up to the BOI. Generally, these models are not concerned with temporal sequence in the meaning of action.

Representatively sampling can be complex. Representatively sampling from extended behavioral interactions in SOI leading to an eventual BOI (e.g., a risky sexual decision) is complicated. Extended social interactions often involve intervening obstacles and challenges, each with decision points impacting the eventual BOI. Ensuring that settings, situations (including extended social interactions) and stimuli are representative for the target audience and target behavior to which we wish to generalize might seem impossible.

Sampling for implementation into DCG. However, we argue that it is not too difficult using the power of enabling technologies such as virtual environments to develop DCG and Experimental conditions built on the DCG base (E_1DCG) (Miller et al., 2019). In order to create a representative sample for the default control group (DCG) in SRD a social scientist takes a series of steps, as provided in Table 2 with an example. This includes identifying the BOI, the POI and identifying and recruiting representative samples of POI, identifying the most frequent settings (MFS) leading up to BOI for POI, extracting details (e.g., cues for POI in MFS), identifying relevant scripts in MFS, identifying the components of those scripts (SCs), and the "entry" and "exit" conditions for each SC, additional details and frequency of response options, and so forth. In the Appendix we elaborate on this process (see also Miller et al., 2019).

Virtual validity check. To insure that a given DCG is representative of everyday life, however, we also use a "virtual validity check" by correlating participant virtual choice to their prior behavior (e.g., past 90 days) in response to similar real-life situations. We found that these coefficients were quite high, approaching values normally associated with test-retest reliability (see for example, Godoy et al., 2013; Smith et al., 2018). Below, we first briefly discuss some of these technologies and their use for a science involving SRD.

SRD Virtual Environments

Definitions and background. Virtual environments are artificial environments in which one's actions determine what happens next in the pursuit of goals (Overmars, 2005). Virtual environments require a combination of software and hardware (e.g., virtual reality (VR)[17], computer hardware).[18] In virtual environment without VR, one can push

[17] VR can consume the user's audio and visual senses, activating haptic responses.
[18] Related technologies include augmented reality (AR) in which virtual objects are overlaid onto one's real-life world and mixed reality (MR) that augments as above but also anchors those virtual objects, making it possible to interact with them in the real-world. (Garon, Boulet, Doironz, Beauliu, & Lalonde, 2016; Tepper, Rudy, Lefkowitz, Weimer, Marks, Stern, & Garfein, 2017; Tokareva, J. (2018). Haptic

or click buttons or move a mouse in making choices for one's character (for example in a digital game): This can be very engaging, affording opportunities for studying complex psychological processes in representative social interactions between one's own agent (that one controls) and other agents. When those other agents are "intelligent agents" within a digital game (using software such as PsychSim (Marsella, Pynadath, & Read, 2004)) whose parameters (e.g., for goals and beliefs) can be set to emulate the variability within and between various groups and cultures, one can have the "feel" of very complex extended social interactions (Marsella, Pynadath, & Read, 2004) in a 3-Dimensional animated environment (for example, see Christensen et al., 2013). Those interactions could be dyadic or involve participants interacting with representative team members or manipulations of same (or other variables).

A virtual reality (VR) headset, on the other hand, obscures all but the virtual environment and enables one to feel one's senses immersed in the virtual environment. With VR, the user moves his or her own body through the virtual environment assuming the identity of an avatar as he or she moves through the virtual environment. For example, looking at a virtual environment without VR of a mechanic working on a vehicle one can see the steps the mechanic takes and even choose them. The same simulation with VR, can enable the user to get the "feel" of the mechanics of the actions as if corporeally immersed in the interaction that offers haptic feedback (e.g., feeling one's hand using a tool or lifting and removing objects). The user can interact with other avatars, but at least at present, the nature of those social interactions tends to be more

---

interfaces may be especially interesting for applications with mixed reality and smartphones with digital/VR games "in the wild" for example. (Lee, Sinclair, Gonzalez-Franco, Ofek, & Holz, 2019) restricted (and typically are quite "clunky") than that which is possible in non-VR virtual environments.

Digital video games[19] afford the user features of a virtual environment experience (e.g., the user can control computer keys, pads, etc. to make choices that affect how the action proceeds; receive feedback). This involves software and hardware (e.g., an electronic gaming device; a laptop computer; a smartphone -- all of which today are likely to have the capacity to control graphic images and a TV or other screen for displaying images). Digital video games can involve VR or not[20]: They are often classified by their game genre (e.g., role-play, active, active adventure, adventure, simulation[21], strategy, etc.)[22], game purpose (e.g., entertainment or serious -- such as to enhance health, education, or training), and type of game platform (e.g., for use on mobile phones, personal computers, iPads, or that can be used cross-platform) (Adams, 2013). For SRD, the video game genre most relevant involves simulations. Simulations [23] are typically designed to emulate real-world scenarios[24]. Generally, SRD would fall

---

[19] The broader category of games includes games that are not video games, such as card games. There are other things described as games (e.g., economic games, such as "prisoner's dilemma") that are associated with game theory, involving the modeling of strategic interactions between rational players (see for example, Myerson, 1991). These literatures, however, are well outside the scope of the current work.
[20] The video game industry, including the much smaller VR video game industry, is a major industry, rapidly overtaking the 125 billion dollar annual mark (Gaudiosi, 2016).
[21] A commercial version is the well known video game simulation (and its many variants) called the "Sims."
[22] As Adams (2013) notes there are also subgenres within each of these genre types (e.g., shooter games are a subgenre of action video games).
[23] Simulations are sometimes classified within the category of digital games and sometimes not, depending upon one's specific definition of "video games" and the specific nature of the simulation in question. For our current purposes -- we are interested for research and behavior change purposes in the capabilities of such software and hardware for implementing the "gist" of representative SOI and BOI for a POI within a virtual world -- when it comes to language pertaining to such distinctions (e.g., when is a simulation not a game) we find ourselves in agreement philosophically with Wittgenstein (2009).
[24] Simulators, for example flight simulators, have been used extensively for decades and found to be highly effective (Hays, Jacobs, Prince, & Salas, 1992)

under the category of a "serious game" because it is not primarily designed for entertainment, but rather as a "test-bed" for science and for developing interventions for behavior change. Diving further into these and other game-related distinctions, however, is well beyond the scope of the current work.

Brief comparison of VR and Non-VR digital games. Given their current status, the choice of whether to use VR or not in research and interventions depends upon the researcher's goals and the nature of the planned intervention. At present, VR would likely be a great choice where the premium on research is on the sensory/physical (e.g., eye-hand coordination; sensory systems and multi-sensory cues or triggers involving visual and auditory channels). This is because, with the increasing sophistication of VR, more granular grasping and bodily precision movements are becoming possible, for example with haptic interfaces that the user wears that provide feedback from the touched object as if it is real (Lee, Sinclair, Gonzalez-Franco, Ofek, & Holz, 2019). In contrast, digital non-VR games, can feature underlying software with intelligent agents that affords new possibilities for understanding social processes and enhancing social and communication interactions and skills (such intelligent agents are not currently available in a VR version). Below, we start with a little of the "bigger picture" of virtual games and their utility and then focus on a much smaller subset of virtual environments designed to be more representative of everyday life.

Overview of virtual environment's applied potential. Digital games, including commercial games designed for entertainment and those not designed primarily for entertainment (i.e., serious games), until recently have mostly not used VR. Existing games have shown promise as interventions, whether designed with that goal in mind or not. As the review by Granic, Lobel, and Engels (2010) makes clear, a wide variety of games, including commercially designed games for entertainment, have potential for producing positive motivational, affective, cognitive, and social effects. For example, playing some shooter games (a subgenre of active games), may dramatically enhance player's spatial skills (Uttal et al., 2013), which, in turn, predict achievements in STEM fields (i.e., science, technology, engineering and math) (Wai, Lubinski, Benbow, & Steiger, 2010). This is quite remarkable, especially since many of these commercial games of various genres (including shooter games) are designed for entertainment, and not for education, per se. Although education scholars may learn quite a bit from studying such entertainment games, and why they may be effective for enhancing learning (Gee, 2003), they are neither designed to provide a test-bed for basic science nor designed for "serious game" applications (e.g., serious games can have the primary goal of assessing or changing behavior). Most games with the primary goal of providing entertainment are not designed to provide a representative environment for assessing or changing behavior.

For those games designed as serious games, there are a number of other meta-analyses and reviews pertaining to their effectiveness (e.g., Papastergiou, 2009; Hieftje et al., 2013).[25] Serious games often try to entertain (and be fun) while also

[25] In studies with commercial games as well as randomized trials with serious games, whether VR or virtual games, the control group is most often either a wait-list control group or it's another game. Each has weaknesses, in that effects in the former may be due to using a game without the intervention components critical; The effects in the later are hard to compare because so many variables differ across games (Granic, et al., 2014). Furthermore, for meta-analyses the control comparison with a comparison game differs considerably across studies whereas with a wait-list control group that control condition is more "constant". SRD addresses this issue by first developing the DCG that can be used as the appropriate control in a randomized controlled trial.

advancing learning (e.g., knowledge, developing skills, enhancing mastery, such as of self-monitoring) by using a variety of underpinning theories (e.g., about observational learning; about message tailoring, about persuasion, social cognitive theory, etc.): For example, they may achieve these dual goals with adventure stories in fictional worlds (Thompson, 2012). Although there are a number of simulations/ game simulations that show good transfer of knowledge and skills (e.g., Gaba, Howard, Fish, Smith, & Sowb, 2001; Lateef, 2010), most of these serious games are not designed to be simulations and/or representative of scenarios/choices in everyday life. Our focus below, although not intended to be a thorough review, provides examples of research using games/simulations (whether with VR or not). The games/simulations reviewed are those that could be considered more representative of the scenarios and challenges of everyday life.

More representative virtual environment designed for research, diagnosis, and treatment. Below, we discuss research from this growing literature using more representative designs that involve virtual reality (VR) and non-VR digital game interactive narrative (IN) environments and how we could leverage these technologies to advance a new paradigm for psychological science.

Virtual reality (VR). Below, we consider VR's use in enabling better basic experimental research. We also consider VR's use in creating better applications for diagnosis and/or treatment (e.g., of a mental health disorder).

Enhancing external validity. Blascovich et al.'s (Blascovich, Loomis, Beall, Swinth, Hoyt, & Bailenson, 2002) pioneering work showed that VR environments could be used to assess whether classic experimental social psychology studies *conducted in the lab* (e.g., social influence in gambling situations that are more naturalistic) generalize to the virtual environment (Blascovich et al., 2002). Some researchers have made the argument that virtual environments take an intermediate ground in the perceived "tradeoff" between internal validity and external validity for research and assessment (Blascovich, et al., 2002; Schonbrodt & Asendorpf, 2011; Johnsen, Raij, Stevens, Lind, & Lok, 2007). This seems a reasonable claim within the context of standard experimental designs, that also nicely gets at the issue of external validity.

But, this approach does not address GEL. It is unclear if individuals' *virtual* behaviors in response to *virtual* situational triggers are similar to their behaviors in comparative situations in their everyday lives. Further, while VR has considerable promise, we believe that its promise goes well beyond a "tradeoff" between internal and external validity (see below).

Is VR effective for diagnosis and/or treatment? A recent review of the more than two dozen review articles of the use of VR demonstrates the potential for VR in diagnosis and treatment (Riva, Wiederhold, & Mantovani, 2019), including the diagnosis and reduction of the effects of PTSD and sexual violence (e.g., Rothbaum, Rizzo, & Difede, 2010; Rizzo & Koenig, 2017; Rizzo et al., 2014). It also demonstrates potential for the diagnosis of obsessive-compulsive symptoms, such as compulsions in response to everyday life (van Bennekom, Kasanmoentalib, de Koning, & Denys, 2017), drug relapse susceptibility (Hone-blanchet, Wensing, & Fecteau, 2014), and craving in smokers (e.g., Bordnick et al., 2005) and alcohol abusers (e.g., Bordnick et al., 2008). VR also shows promising effects in the successful treatment of a range of mental health disorders, including anxiety disorders (for reviews see, Lindner et al., 2017; Botella et al., 2017; Maples-Keller et al., 2017; Cardos, David, & David, 2017; Arroll et al., 2017) and diet related disorders, as well as in pediatric domains (see Riva, Wiederhold, & Mantovani, 2019). Most of these VR diagnostic and treatment programs are based on cues/triggers encountered in the VR world that are representative of those afforded in the real-world.

Why is VR so effective? Riva et al. (2019) note that VR is very effective in promoting long-term behavior change across many domains,[26] more so even than prevailing treatment "gold standards" such as cognitive behavioral therapy. They argue that this is due to: (a) embodied simulations and predictive coding, (b) the meaning of presence, and (c) the instantiation of real-world physical rules within the virtual environment. We would add another contributing factor: VR can afford a kind of interactive narrative (IN) during interaction with it, especially if a therapist is guiding the interaction. We explicate these four factors in the four paragraphs below.

Consistent with work on predictive coding (e.g., Clark, 2013; Friston, 2010), Riva, Wiederhold, and Mantovani argue that, "VR shares with the brain the same basic mechanism: embodied simulations" (2019, p. 88). Body ownership and simulation, involving the identification and integration of internally and externally generated multi-sensory channelled input, enables goal-directed behavior in humans and other species (Botvinick & Cohen, 1998; Van Den Bos & Jeannerod, 2002). Research regarding the Rubber Hand Illusion (RHI) paradigm (Botvinick & Cohen, 1998) aids in

---

[26] Price and Anderson (2007) argued that although VR could facilitate a sense of presence, it did not cause a positive treatment outcome.

the inference that the brain creates these embodied simulations. In the RHI, the researcher simultaneously strokes both a real hidden hand and a rubber visible hand (externally generated cues) causing the participant to perceive body ownership over the rubber "hand" as determined by the participant's response to the researcher striking the rubber hand with a mallet.[27]

If, to effectively regulate one's body in the world, the brain creates and maintains simulations of that body operating in the world -- including various interoceptive, motor, and sensory inputs, and links incoming multi-modal patterns of activation to similar prior multimodal neural activation patterns -- what does the brain do with this information? It is hypothesized that humans use this information in predictive coding (Friston, 2010; Friston & Kiebel, 2011; Clark, 2013), in which past neural patterns (e.g., involving multimodal distributed neuron patterns across diverse brain regions supporting the achievement of a specific goal, concept, or emotion activation) are used to make predictions about what will happen or the meaning of what is happening. When the input reactivates a sufficiently similar pattern of these distributed neurons, the individual experiences the action (e.g., Clark, 2013), concept (Barsalou, 2003), or emotion (Barrett, 2017). In a similar way, VR-embodied simulations may "reactivate multimodal neural

[27] This externally generated RHI paradigm it is argued involves mostly top-down processes while alternative similar paradigms (where the user self-initiates movement) involving body ownership (e.g., moving rubber hand (mRHI), virtual hand illusion (VHI), are more likely to be "actively shaped by processes which allow for continuous comparison between the expected and the actual sensory consequences of the actions...[These additional illusions provide the basis with a motor task using VR to test hypotheses about] whether during goal-oriented tasks body ownership may result from the consistency of forward models" (brackets added for clarity) involving both self-generated and distal multisensory cue integration (Grechuta, Ulysse, Ballester, & Verschure, 2019, p. 1).

networks, which have produced the simulated or expected effect before" (Riva et al., 2019, p. 88).

The feeling of "presence" is an important concept in virtual environment generally (in both VR and non-VR digital games). Presence theory (See Lombard & Jones, 2015 for review) developed, in part, as a concept that accounted for the psychological effects of media technology that embody the user, such as robotics and virtual reality (Steuer, 1992; Sheridan, 1992; Biocca, 1997). Generally, presence has been characterized as a "mental state" (Sheridan, 1992), a "perceptual illusion" (Lombard & Ditton, 1997), and a psychological state (Lee, 2004). Extending the predictive coding hypothesis, Riva and colleagues (2019, p. 88) re-conceptualize the concept of "presence" and argue that "the feeling of presence in a space can be considered as an evolutive tool used to track the difference between the predicted sensations and those that are incoming from the sensory world, both externally and internally." Indeed, the level of presence one experiences using VR may be a function of the degree of similarity between the VR's simulation model of the world and that of the brain.[28] If correct, this suggests that the extraordinary promise of VR extends beyond diagnosis and treatment to deeply understanding psychological processes.

---

[28]Researchers have suggested that we need to be careful in designing representative environments, avoiding and testing for problems like the uncanny valley (e.g., where the agents are so similar to the target person that they activate disbelief) or VR sickness (Cobb, Nichols, Ramsey, & Wilson, 1999). However, although one should test for this in a given population (e.g., Benoit, Guerchouche, Petit, Chapoulie, Maneva, Chaurasia, Drattakis, and Robert, 2015), that as participants are more immersed and experience greater presence in more representative environments, and better leverage emerging technologies for their target populations and behaviors of interest (e.g., Garcia-Betances et al., 2015), that virtual environments may provide closer and closer approximations to everyday behavior in similar everyday affording situations without detrimental effects.

The nature of the physics in VR matters. VR simulations involve some virtual scenes (e.g., with 3-Dimensional models of objects, character agents, and the integration of landmarks) that can be merged into an overall model using professional design software and a means (i.e., the game engine) to calculate the relations between the player and this 3-Dimensional model. What is achieved is the simulation of some real-life physics (e.g., collision-detection, gravity, etc.) and representative human behavior in response to it (e.g., human movements in response to stimuli) along with physical and sensory (visual, auditory) capabilities (Mueller et al., 2012). Even in clunky interactions, this can be enough to create a sense of presence (Regenbrecht, Schubert, & Friedmann, 1998). However, more realistic physics-based interactions, such as physics-based hand-object interactions (Höll, Arth, Overweger, & Lepetit, 2018) have been improving dramatically, leveraging augmented and mixed reality, and likely increasing a sense of presence (Antotsiou, Garcia-Hernando, & Kim, 2018). However, for SRD, more than just physical reality is needed. In theory, social reality in VR could also mimic the quality of physical reality in VR. In practice, VR incorporates social presence often in somewhat indirect or superficial ways (e.g., the presence of a sensory-based (e.g., visual, auditory) scene or one or a group of avatars that evokes a social interaction or the sense of one (i.e., the implied presence of others in a scene that mimics a conference room where the player is expected to speak, designed to evoke social anxiety). Where there are interacting avatars in VR, social interactions to date are often primitive and often clunky.

Interactive narrative (IN) games could be used with VR. IN are systems that allow users to take a role in a narrative story and interact with character agents in an environment (Si, Marsella, & Pynadath, 2010). Broadly, IN occurs in a range of virtual environments. For example, in Blascovich et al.'s (2002) virtual environments using VR, he and his team assessed social influence in gambling situations designed to be very similar to those sequences, that are interactive, in gambling situations he had constructed in laboratory studies years before and that can provide the affordances and cues that can evoke gambling scenarios. Similarly, Rizzo and his colleagues have studied mitigating the effects of post-traumatic stress disorder (PTSD) (Rizzo & Koenig, 2017) and sexual violence (Rizzo et al., 2014) using a VR environment with visuals and sounds within which users experience (or might re-imagine) narratives that trigger their prior trauma with the guidance of a trained therapist. VR with IN-like features have also been used to enhance diagnosis of drug relapse susceptibility based on cues/triggers encountered in the virtual reality (VR) world (Hone-Blanchet, Wensing, & Fecteau, 2014). In IN research and interventions, cues and scenes are often chosen to be more representative of the relevant triggering scenarios to afford relevant cue inter-associations present over time and to be of interest for the POI (e.g., PTSD triggering events).

Non-virtual reality. Representative designs without VR have also been used for a range of interventions. This includes those to enhance health-related communication (e.g., Marsella, Johnson, & LaBore, 2000) and those to reduce risky sex using an interactive video (Read et al., 2006), and intelligent agents (Christensen et al., 2013).

IN games with intelligent agents. "Intelligent agents" are so named because these agents can autonomously determine how the action within an IN proceeds based on how other agents and humans respond and negotiate with other agents and humans to achieve their goals (Marsella et al., 2004). Intelligent agents can achieve these feats using a range of underpinning mechanics, sometimes with quite complex and deep psychological models. For example, PsychSim software agents (Marsella et al., 2004; Miller et al., 2011) have a “theory of mind” about the self and all the other players, including the human user, and these intelligent agents try to pursue goals in the negotiation based on their different parameter settings[29]. Most non-VR IN games have intelligent agents, although some VR IN virtual environments are not driven by intelligent agents (e.g., Rizzo et al., 2014)[30]. Intelligent agents in IN environments have been used in a number of serious games (Ritterfeld, Cody, & Vorderer, 2009), and applications for changing behavior (Miller et al., 2009). Our own socially optimized learning in virtual environments using intelligent technologies (SOLVE-IT) game (Christensen et al., 2013; Miller et al., 2009; Miller et al., 2011) features an extensive series of virtual date scenarios across two game levels designed to reduce risky sexual behavior for young men who have sex with men using the PsychSim Software. The parameters underlying the intelligent agents (e.g., affecting goals and beliefs) drive the choice of agents, which in combination with the human decision-maker, drive how the story proceeds. That is, agent parameters (that can be manipulated) play a large role in the emerging social situation. This nationally disseminated intervention was the first intervention to successfully reduce sexual shame (and to do so for young men who have sex with men) and to show that

[29] Here, the underlying person parameters (goals, beliefs) that might drive behavior of those other actors should be considered, bearing in mind the technology used to implement agents in the game and leveraging its capabilities. For example, we manipulated the parameters of intelligent agents (Marsella et al., 2004), including various goal weights and belief parameters to guide automatic scenarios affording different sexual risk challenges (SOI) leading to our BOI, risky sex, for our POI, as in the real-world.

[30] In theory, VR could involve intelligent “other” agents. However, this is currently not the case.

reduction of sexual shame significantly subsequently reduces risky sex in a longitudinal randomized controlled trial (Christensen et al., 2013). The sophistication of IN games with intelligent agents for modeling complex representative social interactions, and for creating representative interaction partners and interactions, make it particularly promising for modeling and understanding representative social behavior, including detailed social interactions over time (and game levels) in a virtual environment.

Game physics used with IN games with intelligent agents. There are a variety of game platforms that can afford physics as in everyday life. One of the most popular is called Unity, https://unity3d.com/unity/, a 3-Dimensional animated game development platform[31] that has a real-time game engine, or software development environment. Authors on this paper have used Unity for developing SOLVE-IT and many other applications. Physics engines (software) built into these games enable computers to create and tell 3-Dimensional objects how to interact in the digital world, affording an increasingly sophisticated "real-life-like" physics that can have GEL (e.g., where objects have mass and respond to gravity, with drag and angular drag, and can have velocity and respond appropriately when given levels of force and torque are applied, given drag). We can record how objects (and agents, including one's self character) are moving, and responding in the world, and relating to other objects and others (e.g., intelligent agents) over time. Advances in voice recognition and sophisticated game physics, like real-world physics, enables the design of virtual environments that are more representative of

[31] Unity games can be built once and then used across over 20 different platforms including on smartphones, iPads/tablets, computers of every sort; it is used in about half of all games developed since 2005, and is relatively easy to learn and use, and has a real-time game engine -- software development environment (https://unity3d.com/unity/features/multiplatform).

real-life SOI and BOI for our POI, facilitating embodied simulations and a sense of social and physical presence.

Neuroscience Measurement in SRD Using Virtual Environments

A wide range of biobehavioral indicators (e.g., eye tracking software; physiological measurements, such as skin conductance, heart rate, blood pressure) could be used with SRD while individuals are playing a game, whether using VR or not (Miller, Jeong, & Christensen, 2019; Miller et al., 2019; Weibel, Grubel, Zhao, & Schinazi, 2018). A large exception involves the use of fMRI with VR, where VR involves the user's own actual bodily movement: This is because one must remain relatively still in a scanner. To try to provide something akin to a VR experience using fMRI[32], researchers have modified the VR headset so that it can be used in a scanner, but it doesn't afford the same sense of bodily presence that bodily movement affords (for a review see Wiederhold & Weiderhold, 2008). Nevertheless, the use of this modified headset using fMRI appears promising for a range of applications: For example, researchers have examined the effects of alcohol intoxication on virtual driving, both behaviorally and neurally (Calhoun et al., 2005), and the use of the modified headset to study its capacity to reduce pain, both in terms of self-report and also neural patterns (Hoffman, Richards, Coda, Richards, & Sharar, 2003; Hoffman et al., 2004; Parsons, Gaggioli, & Riva, 2017 see also, Bohil, Alicea, & Biocca, 2011 who used EEG in conjunction with fMRI). These effects may be extended into social interactions: Schilbach et al. (2006) found neural evidence that human observers can be socially entrained by virtual characters to

[32] Of course the VR headset is also modified so that it does not contain ferromagnetic metals.

whom they attribute communicative intention. The user experience with modified headsets in a scanner (where the user does not actually bodily move as with regular VR) can involve the sense of moving in a virtual space (Wiederhold & Weiderhold, 2008): This is similar to what a user can experience within a non-VR interactive narrative (IN) game (Christensen et al., 2013).

Using representative virtual IN games in fMRI scanners, one can examine how individuals differ in their *neural circuit* responses, for example, to sexually risky decision points versus conversational decision points (Smith et al., 2018). Neurofeedback fMRI studies today strive to diagnose, train (i.e., enhance self-regulation)[33], and monitor neural responses in contexts where atypical neural connectivity may adversely impact behavior (Robineau et al., 2017). SRD could optimize the potential GEL and causal-inferences possible in such work. Furthermore, because each level of the game could represent, for example, a month (yet be played in minutes), SRD IN games offer a "crunched time" capacity potentially useful in capturing hard to observe patterns of virtual behavior over time (e.g., oscillating dynamic patterns) while also collecting neural patterning data.

Dynamics that Might Otherwise be Hidden

One example of dynamics that might be more readily "observable" in a multilevel game is narcissism, a perplexing personality pattern. Within a person, narcissism seems to include both periods of expressed grandiosity and vulnerability (Pinus, Cain, & Wright, 2014; Coleman, Pincus, & Smyth, 2019). Narcissistic individuals have an inflated sense of self-worth, prioritize their own needs and goals over those of

[33] See Emmert et al. (2016) for a meta-analysis of fMRI self-regulation neurofeedback.

others, and believe that they are entitled to better treatment than that afforded others, and constantly seek admiration and recognition (Krizan & Herlache, 2018). The vulnerability pattern likely emerges when the supports for this inflated self-worth are threatened. In that context, narcissistic individuals' self-regulatory capabilities are readily challenged and impaired (J. D. Miller, Lynam, Hyatt, & Campbell, 2017; Pincus, Roche, & Good, 2015). A variety of theoretical perspectives (Morf, 2006; Krizan & Herlace, 2018; Pincus et al., 2015) argue that grandiosity and vulnerability patterns dynamically feed into and reinforce one another over time (e.g., Giacomin & Jordan, 2014, 2016; Gore & Widiger, 2016; Hyatt et al., 2018; for a review see Coleman et al., 2019). However, it is hard to "observe" this potential oscillation pattern behaviorally and even harder to determine the control parameters (environmental; within-person) that underlie it. As Coleman et al. (2019) note, the specific situational stressors that threaten narcissistic individuals have been hard to discern; the same dynamic (e.g., grandiose behavior and non-responsive or dismissive claims or overwhelming rejection/humiliation producing self-regulatory breakdowns) may operate in an interpersonal conversation over a few minutes or over months or years. Measuring these in-the-moment trigger-response interpersonal dynamics under controlled conditions is quite difficult: SRD in a virtual environment -- in conjunction with computational modeling to anticipate critical control parameters for within-person oscillations for a given individual -- is needed.

The Value of SRD for Psychological Science

SRD requires a great deal of "upfront" work (also see Appendix), but we suggest that the potential payoff is worth it. We argue that SRD will enable psychologists to better address criticisms and advance psychology as a science.

Psychology: Science or Scientific-y?

Is psychology a science or just scientific-y? Outside the field, skepticism over whether psychology is or is not a science is not uncommon (Lilienfeld, 2012: 2017): It is also reflected in the occasional senate bill or congressional vote to strip *National Science Foundation* funding or occasional claims that surface in the public (e.g, newspapers, blogs). Berezow, a microbiologist, for example, argued that psychology was not a science because "psychology often does not meet the five basic requirements for a field to be considered scientifically rigorous: clearly defined terminology, quantifiability, highly controlled experimental conditions, reproducibility and, finally, predictability and testability" (Berezow, July 13, 2012). Not surprisingly, psychologists argue they meet these tests (e.g., Wilson, 2012): We share the frustration. But, it is worth noting that well into the 1900's, prominent physicists viewed only physics as a science[34] (Bernal, 1939).[35] So what changed for biology and chemistry? And is there a lesson for psychology?

What Makes Nobel-Worthy Science?

[34] Rutherford, who won a Nobel Prise Prize in Chemistry was overheard noting that science is either physics or stamp collecting.

[35] The physicist Hoffman more recently noted, "When I was in high school, I loved science and mathematics, but I could never get too excited about biology. It seemed like a lot of tedious memorization and ad hoc theories and appeared to lack the coherence, clarity, and universality of physics. This remained my opinion for many years" (Hoffman, 2012, p. 2).

To answer this question, we looked at what makes for unquestionably good "science". We did so by examining the past decade of Nobel Prize[36] winners in the "hard" sciences, a noble aspiration.

Criteria for Nobel science include innovative methods. Many of the criteria for who/what wins a Nobel Prize in the hard sciences appear to involve innovative methods. These are described in more detail below.

Precisely measure the smallest critical unit. The criterion involves affording observation of the smallest critical unit in dynamic interaction with other critical units, often, in real-time. The goal is not only greater precision but greater accuracy (National Academies of Science, Engineering, and Medicine, 2019) in for example, measuring these key units in interaction with other key units/concepts, or in observing how they operate -- within cells -- at lower levels of scale). In chemistry, for example, Nobel-worthy methods enable insight into intercellular communication via hormone-receiving receptors (Nobel Media AB, 2012) and the development of Super-resolved Fluorescence Microscopy (Nobel Media AB, 2014) to look inside living cells in operation and literally see how molecules interact at the nanoscale level.[37]

Improve experimental capabilities. In Physics, Geim and Novoselov (Nobel Media AB, 2010) found innovative ways to create a 2-Dimensional material, graphene, in which a single layer (one sheet) of atoms were arranged in *hexagon* forms**,** opening up new possibilities for exquisitely controlled experiments of electron behavior. Haroche

[36] There is a Nobel prize in physics, chemistry, and medicine/physiology, but not biology per se. Nor is there a Nobel Prize (officially) in the social sciences.
[37] Note that the level of scale here for this precision and dynamic examination of molecules in their context is at the level of molecules and cells, not whole organisms.

and Wineland (Nobel Media AB, 2012) created groundbreaking experimental methods to better measure and experimentally manipulate individual quantum systems.

Use computational methods to illuminate the complex patterns of critical units. The focus is on understanding complex interactions across scale. For example, the Nobel Prize in Chemistry (Nobel Media AB, 2013) was awarded "for the development of multiscale models for complex chemical systems")[38]. Karplus, Levitt, and Warshel used quantum and classical mechanics and computational tools to calculate complex chemical reactions when new molecules are formed, aiding prediction and hypothesis testing.

Afford manipulation and change. For example, Arnold leveraged evolutionary theory for new methods for protein development -- directed evolution of enzymes -- to solve chemical problems (Nobel Media AB, 2018a). In Physics (Nobel Media 2018b) Ashkin was awarded one for tools--optical tweezers-- that can use laser light to move small particles, and living bacteria, without harming them: He also subsequently used them to investigate "the machinery of life".

Additional criteria. We identified two additional common criteria. The second major Nobel criterion involves the relevance and potential that the innovation had for large social impact that accompanied paradigm shifts. For example, Arnold's work involving directed evolution of enzymes could be used to solve social problems (Nobel Media AB, 2018a). The third major Nobel criterion for discoveries in physics, chemistry,

[38] The 1998 Nobel Prize in Chemistry (Nobel Media AB, 1998) was won for computational tools: Pople for developing computational methods for quantum chemistry and Kohn's involved a density functional approach. The 1999 Nobel Prize in Physics (Nobel Media AB, 1999) involving computational modeling was awarded to Veltman and Hooft in quantum field theory.

and biology involved theory, usually this involved the testing of[39] or advancement of cumulative "big" or "bigger theory." Those "bigger theories" included, in physics, the Standard Model in particle physics (Mann, 2010) and Einstein's Theory regarding gravitational waves (Steinicke, 2005); in chemistry, the Periodic Table (Scerri, 2007); in medicine/physiology, Evolutionary Theory (Darwin, 1859/2002)).

A "hard science" Nobel for a Psychologist? Psychologists or those trained in psychology have been awarded a Nobel Prize in Medicine/Physiology, most recently in 2014 for discovering the brain's neural positioning system, O'Keefe (a psychologist/ neuroscientist) won his portion for the discovery of "place cells" (CA1 hippocampus area (see O'Keefe & Nadel, 1978)) and the Mosers, both neuroscientists, won theirs for discovering a correspondence between a *hexagonal* grid with evenly spaced (and same direction/size) electrode spike firing of nerve cells in the rat's brain in the dorsocaudal medial entorhinal cortex (dMEC) during the rat's movement/positioning in his environment that coordinated with the CA1 hippocampus area "place cells" (Moser, Kropff, & Moser, 2008). Drawing from the nobel criteria above, this work identified *smaller critical units* (i.e., grid cells) that in interaction with one another and with other, place cells, provide insight into fundamental questions *across scale* about how brains have the capacity to navigate in our environment (e.g., representing position, direction, and velocity). Additional discoveries since this Nobel was awarded indicate that the uniform hexagon space grid "warps" in line with the reward learning histories for specific motivations/goals in a given context (Butler, Hardcastle, & Giocomo, 2019; see also

[39] Providing the basis for refuting "bigger theory" or strong systematic evidence that undermines major theoretical assumptions of established theory could also provide a basis for an award in science at this level.

work by Boccara, Nardin, Stella, O'Neill, & Csicsvari, 2019).[40] The content of this work and its innovative and *paradigm changing methods* seem tantalizingly within our grasp as psychologists. [41]

Aspirational Psychological Science: SRD's Role

Using the roadmap provided by Nobel Prize awards as a guide, what is the science towards which psychologists should aspire? SRD could help move us towards an aspirational science, advancing each of the following:

Shared definitions of units of interest. For social and personality psychologists, key units include, for example, situations (or contexts), for which we do not have shared definitions. Default control groups (DCGs), provide an initial concrete implementation of a person-in-context model (e.g., like a model system in biology) of one or more situations involving ongoing social interactions, and physical affordances. With feedback (via virtual validity checks) to assure GEL, DCG could provide cumulatively more precise and accurate shared definitions of contexts and person-in-context interactions over time.[42]

---

[40] Furthermore, additional researchers, building on this work, investigating the combination of multisensory self-motion and place/landmark information ***in virtual environments with mice*** developed a network model whose principles were further tested, moving scientists towards a theoretical framework for understanding how environment and self-cues produce the spatial representations guiding goal-directed behavior (Campbell, Ocko, Mallory, Low, Ganguili & Giocomo, 2018).

[41] Indeed, the hippocampus plays a significant role across rodents and humans in decision-making involving approach-avoidance conflict: it is key, however, to study these motives concurrently (Bach, Guitart-Masip, Packard, Miró, Falip, Fuentemilla, & Dolan, 2014); Ito & Lee, 2016; Oehrn, Baumann, Fell, Lee, Kessler, Habel, Hanslmayr, & Axmacher, 2015; O'Neil, Newsome, Li, Thavabalasingam, Ito, & Lee (2015).) reminding us of N. E. Miller's (1944) classic approach-avoid conflict research (and the importance of measurements in the rat's *movement in space* as it negotiated this conflict). This suggests the need to revisit this work on movement to assess this conflict (Boyd, Robinson, & Fetterman, 2011) using today's technologies (e.g., Oculus Rift/VR; animated characters interacting with humans) similar to what has and is currently being done, involving fine-grained head movements in both approach and avoid motivations in conflict situations Jeong, Feng, Krämer, Miller, & Marsella (2017).

[42] In building initial DCG, for example, we are concurrently testing assumptions about key features in it (e.g., settings and their affordances; structures in it (e.g., scripts); beginning and ending points, etc.). These evidence-based assumptions, for example, can be challenged (e.g., with comparisons with alternative models; by experimentally eliminating/altering aspects of the model in experimental comparisons to judge

Precision measurement of critical units in dynamic interaction over time. It can do so with automatically recorded precise observations (e.g., virtual choices; the physics of movements, such as avoiding/approaching others/objects) of an individual's agent behavior interacting within social interactions and contexts; sophisticated intelligent social agents with known representative underpinning parameters; and with the capacity of SRD in virtual environments to "crunch time," this could enable precision examination of the complex triggers and oscillating behavior patterns that can emerge for individuals in interaction with others over long periods of time.[43]

Concurrent measurement of underpinning brain patterning. Concurrent measurement of underlying mechanisms while engaged in representative everyday interactions during extended social interaction is possible today (e.g., with fMRI, for example, Smith et al., 2018).

Experimental manipulation, causal inference, and change assessment. It is possible with SRD to compare an experimental group to a control group that differs only in the independent variable of interest. The control group itself has generalizability to everyday life (Miller et al., 2019).

Use of computational models and modeling "experiments." Computational models (Read et al., 2010; Read, Smith, et al., 2017; Read, Droutman, Smith, & Miller, 2017) were used in addressing prior puzzles in personality and social psychology such

their altered virtual validity) to enhance cumulative science precision, accuracy, and insight into when, why, and how they differ in terms of impact.

[43] Furthermore, virtual validity checks in real-time (e.g., using smartphone and sensor technologies, including ecological momentary assessments (Shiffman, Stone, & Hufford, 2008) afford continued feedback and opportunities for cumulative measurement and prediction improvement over time.

as how there could be more within-person variability across situations than between-person variability and still have stable traits such as the "Big-5".[44] Such computational experiments can guide SRD[45] development (Miller, Jeong, & Christensen, 2019) because they suggest, for example, the need to measure certain affordances in situations in everyday life and insure that those are represented in virtual environments as in everyday life for our BOI in our SOI and for our POI.[46] [47] Data afforded in SRD designed with these computational models in mind can be used to further test these models computationally (i.e., do we get the same results from a computational modeling of the same features as we do in terms of participant behaviors in virtual environments with SRD). This process can generate new hypotheses for testing in SRD.

Translation methods to optimize scalability[48] and broader societal impact. SRD provides one possible solution to better methods and more rapid translation for

---

[44] In building their computational models, Read and his colleagues argued that humans have universal approach and avoid systems and nested in them, universal goals: But the relative levels of chronic goals differ between individuals. Situations have different goal affordances as well. As individuals move into different situations (e.g., a friend appears; an alarm goes off), the situational affordances change: These combine with chronic goal activations to affect current competing goal activations, with the most activated goal driving behavior. Computational models virtual personalities (VP) -- where VP chronic activations were systematically manipulated -- indicated that there was tremendous within-person variability in behavior across situations, but at the same time entering each VP's data (as we would for real subjects), and performing factor analyses produced across persons, the "Big 5".

[45] In a way a given SRD DCG could be our "best guess" instantiation of the probability distributions of cues and sequences that constitute a specific context and sequential options and consequences in the real-world. As suggested earlier, this seems analogous to the "model system" concept so critical in modern biology.

[46] In addition, because computational models can be used across scale (e.g., the interpersonal level, the individual level, and the neural level) to address personality and social psychological dynamics in producing emergent behavior (e.g., Read, Brown, Wang, & Miller, 2018), they can also suggest (across scale, for example in fMRI studies) what to measure and afford in building SRD.

[47] In short, computational modeling, since it requires the math and precision to build, provides psychologists with new methods in our toolkit for illuminating hidden assumptions and theoretical gaps, while also affording ways to iteratively build and improve SRD as well as testable theories (Marsella & Gratch, 2016; Marsella, Gratch, & Petta, 2010; Vallacher, Read, & Nowak, 2017; Farrell & Lewandowsky, 2018)

[48] It can take a decade or more for basic science to produce useful applications (Morris, Wooding, & Grant, 2011).

broader impact. That is, if found effective, some of that experimental work could afford interventions that could rapidly move from experimental lab efficiency to effectiveness trials with broad utilization on a national level with relatively low cost to the public on a *per capita* basis (Christensen et al., 2013). Some game platforms (e.g., Unity) such as the one we used for our SOLVE game (Christensen et al., 2013), are highly cross-platform capable -- for example, game interventions developed for computers, can be easily implemented on other platforms (e.g., smartphones). These games could be extended as smartphone interventions or with other smartphone interventions, such as just-in-time adaptive interventions (JITAI), into individuals' everyday life[49] (Nahum-Shani, Hekler, & Spruijt-Metz, 2015; Nahum-Shani et al., 2018; Spruijt-Metz et al., 2015)[50]. Indeed, a recent meta-analysis suggests that JITAI can be quite effective (Wang & Miller, in press).

Reach towards "bigger theory". So many of our theories in social and personality psychology -- from the perspective of hard scientists -- would probably be viewed as "mini theories" where it's hard to see how it all "adds up". Work in these, and most areas of psychology, is generally not tethered to and integrated into a "bigger theory". We elaborate on the "bigger theory" issue in its own section below and how

[49] Technological advances here are rapid, including in exquisite capabilities for voice recognition and emotion differentiation (see for example, Huang & Narayanan, 2017; Somandepalli et al., 2016) and the capacity to "pick up" complex contextual cue reactivity in craving (Traylor, Parrish, Copp, & Bordnick, 2011).

[50] In this era of "big data" (Cai & Zhu, 2015; Kitchin, 2014; Provost & Fawcett, 2013), one question is what will we do with so much rich and complex data? Machine learning may provide one set of answers, but the cues, and the relationships among them that go into these algorithms can often be a "black box." Furthermore, these cues may or may not be the cues that humans use in the same way (Cai &Zhu, 2015). SRD is a methodology through which big data can be leveraged to better create systematic control and experimental groups and to more systematically test how to "structure" data contextually to build better predictive models of human smart-phone and sensor data patterns over time.

SRD could also assist there. In all of these ways, we argue, that SRD would concurrently advance psychology as a science, and therefore, our place within the "hard" sciences.

"Bigger Theory"and SRD

Given emerging findings in cognitive science and neuroscience, we believe that psychology, broadly, may be "on the cusp" of an exciting paradigm shift if we can examine and manipulate *features in context* that are important and representative of those that humans encounter (and are differentially motivated by) in their real-world ecologies. Since many of the authors are personality-social psychologists or from related social sciences, we use examples of social concepts to suggest how we could do this.

Bigger Theory Candidate: Predictive Coding

Hierarchical prediction. Making sense of and acting in the world is what we do all the time, whether in virtual environments or in everyday life. It is nonetheless complex and not direct. Predictive coding is a major theory of how the brain is adapted to make probabilistic inferences (Clark, 2013; Friston & Kiebel, 2011).[51] Predictive coding affords a universal explanatory principle for the operation of the human brain and mind (Bar, 2011a; Friston & Kiebel, 2011). As such, it's a promising candidate for "bigger theory" in psychology -- as biologists, chemists, and physicists might think of it -- that also could guide us in developing systematic representative environments to better understand individuals in their interaction with one another in context. Such SRD designed virtual environments could also guide our thinking about predictive coding in

[51] Although Clark (2013) mostly presents one predictive coding algorithm, different predictive coding models using alternative algorithms still vye for which better capture the data and which is the most neurobiologically plausible (Spratling, 2017).

understanding social construction of causal meaning and social behavior in representative, dynamic, and contextualized person and situation interactions over time.

Clark (2013), reviewing work in cognitive science, computational modeling, and neuroscience, argues for predictive coding: That is, that the brain uses a hierarchical, multi-modal (e.g., Mesulam, 1998; Rauschecker & Tian, 2000; DeWitt & Rauschecker, 2012) "bidirectional cascade of cortical processing," generatively, to minimize prediction error between "top-down expectations or predictions" and "incoming sensory inputs" (p. 181).[52] In the domain of perception -- and also in the domain of action -- humans in trying to track a visually presented scene use prior knowledge, and "top-down" knowledge, to generate "a kind of 'virtual version' of the sensory data" (Clark, 2013, p. 182). For example, imagine a social outcome or consequence: For example, Mary was just injured. We might ask, "What or Who caused this?" "Why?"

As Friston notes, starting with the consequence or effect, the brain essentially works backwards to identify the cause. In doing so, we use causal structures (presumably evolved) in the brain "that distil the causal regularities in the sensorium and embody them in models of their world" that we can use to predict consequences (Friston, 2013, p. 212-213). The task is complex because there are many potential causes (e.g., of Mary's injured state). What might be those structures (each represented here with an arbitrary letter) that our brains use to make causal inference about what happened here.

$$[A\ D\ G\ I\ X\ T\ P\ N] \quad \rightarrow \text{With what consequence/effect}$$

---

[52] Friston (2013, p. 212) notes that "predictive coding is a consequence of surprise minimisation, not its cause."

Read and Miller (1998) argued that each of these "structures" or "slots" are likely to reflect "universal" linguistic concepts[53] across cultures (Wiezbicka, 1992) that include "want", "as well as all of the words in the following 'story': I want this, you do this, this happened, this person did something bad, and something bad happened because of this." (Read & Miller, 1998, p. 49). And, as Friston (2013) notes, the neuroscience literature supports at least two likely candidates, those associated with "*separable attributes of 'what' and 'where' [translating] into separate neuronal representations in segregated visual pathways"* (Friston, 2013, p. 212-213, brackets added).[54] Furthermore, different orderings of the same behaviors produce remarkably different social causal inferences (Read, Druian, & Miller, 1989): That suggests, "when" behavior relative to other behavior occurs, matters[55]. Thus, candidates for the causal "slots" for making causal inferences about the consequences/effects (e.g., Mary's injury) that occur in social interactions may include the following candidates in brackets:

[Who[56]/what] [Did/Said this] [to Whom/What] [How][Where][When]→ effect[57,58] [Why]

---

[53] Read and Miller (1998) also closely examined the developmental literature (Read & Miller, 1995). For example, young children have a readiness to communicate wants (Gelman, 1990).

[54] The frontal cortex plays a domain general cognitive control function, selecting among competing representations and shifting and weighting algorithms between dorsal and ventral multimodal streams and numerous points of integration across ventral and dorsal streams (Bornkessel-Schlesewky et al., 2015a, 2015b.)

[55] A critical feature in experiments is timing, the independent variable for example must precede the dependent variable as one important criterion for causal inference. Of course, in everyday life, individuals use many of the criteria we use in experiments to make their own everyday causal inferences about the meaning of sequences of behavior.

[56] Face processing (and the anticipation of face processing) is especially associated with the fusiform face area (FFA) of the brain (Furl, Garrido, Dolan, Driver, & Duchaine, 2011). That is, specific "who" or "whom" assessments in understanding sequences of actions may be based on connections in one's representation there.

[57] Roseman (2011) provides a hierarchical motive-based model pertaining to emotional "effects" that may serve to also motivate (e-motion) action. Might this theoretical model suggest possible neural (perhaps narratively based) slot unit linked underpinnings?

[58] Exciting work (i.e., Chang, Gianaros, Manuck, Krishnan & Wagner, 2015) in multivoxel pattern analysis (MVPA) that appears to afford sensitive and specific neural signatures for affect induced stimuli (e.g.,

And, the concrete narrative that might be generated in a situation as we tried to understand why Mary was injured, using these slots, might be the following:

John shoved Mary hard at the luggage carousel just now causing her injury:

He was in a hurry and didn't care if he hurt her in the process.

Just as only a few letters (26 in English) afford thousands of words, these "causal slots" of a scenario (we call a "plot unit") can afford an almost infinite number of concrete scenario descriptions, some so recurrent that in a given culture we may give them an economical conceptual name, such as here describing John's behavior as aggressive or using trait terms if John has done things like this repeatedly (e.g., aggressive): Indeed, consistent with Read and Miller (1998), underlying many social concepts may be neurally linked "slots" in one or more of these "plot units" (see Figure 3). Because there are so many ways that each of these categories of slots can be filled, there are many alternative causal inferences that could be activated across persons in understanding the meaning of a given sequence of behavior (although this set of alternatives is not unlimited). As humans move from scene to scene, top-down inferences about a former scene can guide the meaning of a new or upcoming scene, dynamically enabling the construction of the interpretation that best "explains" what is happening, and reduces errors of prediction (due to surprise, Friston, 2013). If these are important slots in making causal inferences, what's the hierarchical nature of this process?

---

aversive images) could benefit from fMRI recording during participant SRD representative scenarios engagement. It is an intriguing possibility that we could examine if such stable neural specific signature activations and their links recapitulate nongoing narratives (and conceptual plot units) in social interaction.

Concretizing hierarchical prediction in social interaction. Read and Miller (1998) used a recurrent neural network interactive activation and competition (IAC) modeling approach (McClelland & Elman, 1986; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982): IAC is one approach, to help concretize the prediction processes humans may use to infer causal meaning in social interaction. Like the brain itself, an IAC approach assumes hierarchical structures and a parallel constraint satisfaction process in making causal inferences regarding the meaning of actions. As illustrated in Figures 3 and 4 (see Read & Miller for further detail), major social concepts (e.g., traits, states, relationships, roles, beliefs/attitudes, and so forth)[59] may "top down" guide our causal interpretation of the ongoing social interaction at the event and sequence levels. Bottom up processes (e.g., from the feature analysis[60] and identification levels) might send "error signals" up, producing surprise (e.g., Clark, 2013; Friston & Kiebel, 2011) and the need for adjustments in predictive coding (Friston, 2013). Figure 4 helps illustrate some of the links, and spreading activation involved in settling on a competing interpretation of an ongoing social interaction and how it changes with new input.

Event models are predictive models of the near future that provide a "top-down" frame or bias that allows the person perceiver to fill in and disambiguate ambiguous information. Consistent with the idea of events as predictive models, an event typically

---

[59] These social concepts themselves are apt to be based on underlying learning histories with respect to various combinations of plot units.

[60] There are extensive literatures focused on "person"/"object" perception as well as action perception-- and the links among these to social judgments. For example, there is considerable work on features underpinning face perception and the relationship of these features to social judgements such as dominance and competence (e.g., Todorov, Dotsch, Porter, Oosterhof, Falvello, 2013) or attractiveness (e.g., Bronstaad, Langlois, & Russell, 2008; Todorov et al., 2013).

ends when there is high prediction error; at this point the perceiver starts to build a new event model, and when possible, integrate it with the prior event (Radvansky & Zacks, 2014). As perceivers may create hierarchies of event models at different "grain" sizes, prediction error and likelihood of updating a model may differ at different "grain" sizes.

Predictive Coding in Building Virtual Environments for SRD. The literature on predictive coding makes it clear that *context matters tremendously*. For the brain, prior knowledge, including what just happened and what we believe the perceiver/doer is anticipating/expecting matters (e.g., Clark, 2013).[61] *It matters in experiments.* As Clark (2013) notes, this "means that we need to be very careful when generalizing from ecologically strange laboratory conditions [e.g., faces without bodies] that effectively deprive us of such ongoing context." (p. 203, material in brackets added).[62] This is a major reason why our laboratory contexts need to be representative of everyday life: If not, it's neither clear what our findings really mean nor whether our findings are likely to have GEL. And *that,* we would argue, is critical for our status as a science.

How does predictive coding help guide SRD development? In building a virtual environment for SRD based on predictive coding, how do we start? Predictive coding suggests that in developing DCG that are representative for our purposes, we should insure in our virtual environment that game players can use the same process (and can

[61] As Clark notes, Helmholtz (1860) had "the key idea that sensory systems are in the tricky business of inferring sensory causes from their bodily effects. This in turn involves computing multiple probability distributions, since a single such effect will be consistent with many different sets of causes distinguished only by their relative (and *context dependent*) probability of occurrence" (Clark, 2013, p. 182, italics added).

[62] For other discussions pertaining to contextualization and its importance, see Kveraga et al., (2007), Bar (2007), Barrett and Bar (2009), and Fabre-Thorpe (2011).

access the same representative features) for meaning construction during an unfolding social interaction scenario in a virtual IN game as that same POI in the SOI for the BOI would need in everyday life.

We know a great deal about the features underpinning scenario construction, and in expanding on that and leveraging it in creating SRD, we can systematically (and with recurrent feedback from participants in developing our SRD to achieve this) reproduce in our virtual worlds what humans need (e.g., in the way of cues) to make inferences they are likely to make in these same everyday contexts. We can test hypotheses and build theories about probability distributions in these representative social virtual worlds that are more consistent with those in the real-world while insuring our science fits with work emerging across psychology (e.g., cognition, decision-making, language, memory, perception, social interaction, speech) as well as in the biological and neuroscience literatures.[63]

Advancing Psychological Science

SRD as a Tool for "Crisis Management"

As the above discussion suggests there are a number of benefits to SRD, including its potential to grapple with current and recurring crises in the field and to leverage criticisms to develop new methods to address those criticisms and advance our science. This includes the relevance crisis (Pettigrew, 2018), and building capacity to leverage theory and research for effective, impactful interventions. It also includes the related

[63] Computational models that take priors (prior constructions; top down concepts) into account in guiding subsequent causal inferences in ongoing social interactions. This could help model and predict users' causal meaning inferences in interacting in virtual (and to the extent possible real-life corresponding) situations. Those computational models could also generate testable hypotheses for participants causal inferences within and about both virtual and real-life ongoing situations over time.

generalizability of our work to everyday life (GEL), the crisis in homogenous samples of POI (Henrich et al., 2010), and the replication crisis (National Academies of Science, Engineering, and Medicine, 2019). Other emerging crises are the need for paradigms that better bridge levels of scale (Cacioppo et al, 2000), for example, between behavior in extended social interaction and neural circuit patterning, and finally, the capacity to build a cumulative science, involving increasingly more precise methods/measures and an understanding of how these experiments fit in the fabric of a "bigger theory," affording prediction and testability (see above).

Integration of Approaches is Essential

Cronbach's (1957) APA presidential address, published in the *American Psychologist,* noted that there were historically two streams of research differing in method and approach: experimental and correlational. He argued that integrating these approaches was critical: Otherwise, he said,

> they can give only wrong answers or no answers at all regarding certain important problems…A united discipline will ... be concerned with the otherwise neglected interactions between organismic and treatment variables. Our job is to invent constructs and to form a network of laws which permits prediction. From observations we must infer a psychological description of the situation and the present state of the organism. Our laws should permit us to predict from this description, the behavior of organism-in-situation (pp. 681-682).

The current approach we have proposed suggests ways to move further towards this goal. We argue that we may be at the "tipping point" for a paradigm shift in experimental design. Systematic Representative Design, as envisioned here, combines correlational and experimental approaches, and affords generalizability to everyday life as well as the capacity for experimental causal inference.[64] A technology-enabled SRD could enhance our science in myriad ways (e.g., precision, robustness and reliability, generalizability to everyday life, cumulative potential for bigger theory, usefulness) and fill the context gap. Systematic Representative Design could also help bridge historical divides, provide tools across scale, support new persons-in-situation methods, better interface with cognitive science, neuroscience, computational science, and artificial intelligence, and help better claim psychology's place in the "hard" sciences.

---

[64] Given a predictive coding approach to developing SRD, there is also the possibility of building in representative environments that afford opportunities for examining variability in how individuals make causal inferences within, for example, a given DCG and in assessing what can alter those patterns (e.g., with experimental groups built on the DCG base). .

References

Abbott, V., & Black, J. B. (1986). Goal-related inferences in comprehension. In J. A. Galambos, R. P. Abelson, & J. B. Black (Eds.), *Knowledge structures* (pp. 123–142). Hillsdale, NJ: Erlbaum.

Adams, Ernest (2013). *Fundamentals of game design* (3rd ed.). San Francisco: New Riders. ISBN 0-321-92967-5.

Allport, G. W. (1968). The historical background of social psychology. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., Vol. 1, pp. 1–80). Oxford, England: Addison-Wesley.

Antotsiou, D., Garcia-Hernando, G., & Kim, T. K. (2018). Task-oriented hand motion retargeting for dexterous manipulation imitation. In *Proceedings of the European Conference on Computer Vision (ECCV)*(pp. 0-0).

Araújo, D., Davids, K., & Passos, P. (2007). Ecological validity, representative design, and correspondence between experimental task constraints and behavioral setting: Comment on Rogers, Kadar, and Costall (2005). *Ecological Psychology, 19*, 69–78. http://dx.doi.org/10.1080/10407410709336951

Argyle, M., Furnham, A., & Graham, J. A. (1981). *Social situations.* Cambridge, England: Cambridge University Press. Retrieved from http://www.cambridge.org/ca/academic/subjects/psychology/social-psychology/social-situations?format=PB

Aronson, E., Ellsworth, P. C., Carlsmith, J. M., & Gonzales, M. H. (1990). *Methods of research in social psychology*, 2nd ed. New York, NY: McGraw-Hill Publishing.

Arroll, B., Wallace, H. B., Mount, V., Humm, S. P., & Kingsford, D. W. (2017). A systematic review and meta-analysis of treatments for acrophobia. *Medical Journal of Australia*, *206*(6), 263-267.

Bach, D. R., Guitart-Masip, M., Packard, P. A., Miró, J., Falip, M., Fuentemilla, L., & Dolan, R. J. (2014). Human hippocampus arbitrates approach-avoidance conflict. *Current Biology*, *24*(5), 541-547.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences 11*(7), 280–89.

Bar, M. (Ed.). (2011a). *Predictions in the brain: Using our past to generate a future*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195395518.001.0001

Bar, M. (2011b). The proactive brain. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 13–26). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195395518.003.0010

Barker, R. G. (1963). The stream of behavior as an empirical problem. In R. G. Barker (Ed.), *The stream of behavior: Explorations of its structure & content* (pp. 1–22). East Norwalk, CT: Appleton-Century-Crofts. http://dx.doi.org/10.1037/11177-001

Barker, R. G., & Wright, H. F. (1955). *Midwest and its children: the psychological ecology of an American town*. Oxford, England: Row, Peterson.

Baron, R., & Boudreau, L. (1987). An ecological perspective on integrating personality and social psychology. *Journal of Personality and Social Psychology, 53*, 1222-1228.

Baron, R. A., Byrne, D., & Suls, J. (1989). Attitudes: Evaluating the social world. In R. A. Baron, D. Byrne, and J. Suls., *Social Psychology, 3rd Ed. MA: Allyn and Bacon, 709-101.*

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1-23.

Barrett, L. F. & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1521), 1325–34.

Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, *18* (5-6), 513-562.

Beans, C. (2018). News feature: What happens when lab animals go wild. *Proceedings of the National Academy of Sciences, 115*(13), 3196–3199. http://dx.doi.org/10.1073/pnas.1803284115

Benoit, M., Guerchouche, R., Petit, P.D., Capoulie, E., Manera, V., Chaurasia, G., Drettakis, G., & Robert, P. (2015). Is it possible to use highly realistic virtual reality in the elderly? A feasibility study with image-based rendering. *Neuropsychiatric Disease and Treatment, 11*, 557-563. doi: 10.2147/NDT.S73179

Berezow, A. B. (2012, July 13). Why psychology isn't science. *LA Times*. Retrieved May 20, 2019 from https://www.latimes.com/opinion/la-xpm-2012-jul-13-la-ol-blowback-pscyhology-science-20120713-story.html

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist, 37*, 245–257. http://dx.doi.org/10.1037/0003-066X.37.3.245

Bernal, J. D. (1939). *The social function of science.* Abingdon, UK: G. Routledge & Sons Limited.

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-mediated Communication*, *3*(2), JCMC324.

Black, J. B., Galambos, J. A., & Read, S. J. (1984). Comprehending stories and social situations. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 45–86). Hillsdale, NJ: Erlbaum.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*, 103–124. http://dx.doi.org/10.1207/S15327965PLI1302

Boccara, C. N., Naedin, M., Stella, F., O’Neill, J., & Csicsvari, J. (2019). The entorhinal cognitive map is attracted to goals. *Science, 363*(6434), 1443-1447. DOI: 10.1126/science.aav4837

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, *12*(12), 752.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234. http://dx.doi.org/10.1207/s15327957pspr1003_2

Bordnick, P. S., Graap, K. M., Copp, H. L., Brooks, J., & Ferrer, M. (2005). Virtual reality cue reactivity assessment in cigarette smokers. *CyberPsychology & Behavior*, *8*(5), 487-492.

Bordnick, P. S., Traylor, A., Copp, H. L., Graap, K. M., Carter, B., Ferrer, M., & Walton, A. P. (2008). Assessing reactivity to virtual reality alcohol based cues. *Addictive behaviors*, *33*(6), 743-756.

Bornkessel-Schlesewsky, I., Schlesewsky, M., Small, S. L., & Rauschecker, J. P. (2015a). Neurobiological roots of language in primate audition: common computational properties. *Trends in Cognitive Sciences*, *19*(3), 142–150. http://doi.org/10.1016/j.tics.2014.12.008

Bornkessel-Schlesewsky, I., Schlesewsky, M., Small, S. L., & Rauschecker, J. P. (2015b). Response to Skeide and Friederici: the myth of the uniquely human "direct" dorsal pathway. *Trends in Cognitive Science. 19*(9), 484-485. DOI: 10.1016/j.tics.2015.05.010

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent progress in virtual reality exposure therapy for phobias: a systematic review. *Current Psychiatry Reports*, *19*(7), 42.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*(6669), 756.

Boyd, R. L., Robinson, M. D., & Fetterman, A. K. (2011). Miller (1944) revisited: Movement times in relation to approach and avoidance conflicts. *Journal of Experimental Social Psychology*, *47*(6), 1192-1197.

Brewer, M. B., & Crano, W. D. (2014). Research design and issues of validity. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 11–26). New York, NY: Cambridge University Press.

Bronstad, P. M., Langlois, J. H., & Russell, R. (2008). Computational models of facial attractiveness judgments. *Perception, 37*(1), 126-142. doi:http://dx.doi.org.libproxy1.usc.edu/10.1068/p5805

Bruner, J. (1987). Life as narrative. *Social Research, 54*(1), 11–32. Retrieved from https://www.jstor.org/stable/40970444

Bruner, J. (1990). *The Jerusalem-Harvard lectures: Acts of meaning*. Cambridge, MA: Harvard University Press.

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review, 50*, 255–272. http://dx.doi.org/10.1037/h0060889

Brunswik, E. (1944). Distal focussing of perception: Size-constancy in a representative sample of situations. *Psychological Monographs, 56*, i–49. http://dx.doi.org/10.1037/h0093505

Brunswik, E. (1947). *Systematic and representative design of psychological experiments; with results in physical and social perception*. Oxford, England: U. of California Press.

Brunswik, E. (1952). *The conceptual framework of psychology. (Int. Encycl. unified Sci., v. 1, no. 10.)*. Oxford, England: Univ. Chicago Press.

Brunswik, E. (1955a). In defense of probabilistic functionalism: A reply. *Psychological Review, 62*, 236–242. http://dx.doi.org/10.1037/h0040198

Brunswik, E. (1955b). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193–217. http://dx.doi.org/10.1037/h0047470

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.

Butler, W. N., Hardcastle, K., & Giocomo, L. M. (2019). Remembered reward locations restructure entorhinal spatial maps. *Science*, 363 (6434): 1447-52. DOI 10.1126/science.aav5297.

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin, 126*, 829–843. http://dx.doi.org/10.1037/0033-2909.126.6.829

Cai, L. and Zhu, Y., 2015. The challenges of data quality and data quality assessment in the Big Data era. *Data Science Journal*, 14, p.2. DOI: http://doi.org/10.5334/dsj-2015-002

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1983). Beyond external validity. *Journal of Consumer Research, 10*(1), 112–114. Retrieved from http://doi.org/10.1086/208950

Calhoun, V. D., Caralho, K., Astur, R., Pearlson, G. D. (2005). Using virtual reality to study intoxication effects on the neural correlates of simulated driving. *Applied Psychophysiology and Biofeedback*, *30*(3), 285-306.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54*, 297–312. http://dx.doi.org/10.1037/h0040950

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton, Mifflin and Company. (Original work published 1963)

Campbell, M. G., Ocko, S. A., Mallory, C. S., Low, I. I., Ganguli, S., & Giocomo, L. M. (2018). Principles governing the integration of landmark and self-motion cues in entorhinal cortical codes for navigation. *Nature Neuroscience.* PubMedID 30038279

Cardoş, R. A., David, O. A., & David, D. O. (2017). Virtual reality exposure therapy in flight anxiety: A quantitative meta-analysis. *Computers in Human Behavior*, *72*, 371-380.

Ceci, S. J., Kahan, D. M., & Braman, D. (2010). The WEIRD are even weirder than you think: Diversifying contexts is as important as diversifying samples. *Behavioral and Brain Sciences, 33*, 87–88. http://dx.doi.org/10.1017/s0140525x10000063

Chang, L. J., Gianaros, P. J., Manuck, S. B., Krishnan, A., & Wager, T. D. (2015). A

Sensitive and Specific Neural Signature for Picture-Induced Negative Affect. *PLoS Biol, 13(*6), e1002180.

Christensen, J. L., Miller, L. C., Appleby, P. R., Godoy, C. G., Marsella, S. C., & Read, S. J. (2013). Reducing shame in a game that predicts HIV risk reduction for young adult MSM: A randomized trial delivered nationally over the web. *Journal of the International AIDS Society, 16 (Suppl 2)*, 1–8. Retrieved from http://www.jiasociety.org/index.php/jias/article/view/18716 | http://dx.doi.org/10.7448/IAS.16.3.18716

Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science, 4*, 5–6. http://dx.doi.org/10.1111/j.1745-6924.2009.01091.x

Clark A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral And Brain Sciences, 36*, 181–204. doi:10.1017/S0140525X12000477

Cobb S, Nichols S, Ramsey A, Wilson J. (1999). Virtual reality-induced symptoms and effects. *Presence,* 8, 169–186.

Coleman, S., Pincus, A., & Smyth, J. (2019). Narcissism and stress-reactivity: A biobehavioural health perspective. *Health Psychology Review.*, *13*(1), 35–72. https://doi.org/10.1080/17437199.2018.1547118

Cornwell, B. (2015). *Social sequence analysis: Methods and applications* (Vol. 37). Cambridge, England: Cambridge University Press.

Costabile, K. A., Shedlosky-Shoemaker, R., & Austin, A. B. (2018). Universal stories: How narratives satisfy core motives. *Self and Identity, 17*, 418–431. http://dx.doi.org/10.1080/15298868.2017.1413008

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*, 671–684. http://dx.doi.org/10.1037/h0043943

D'Argembeau, A., Cassol, H., Phillips, C., Balteau, E., Salmon, E., & Van der Linden, M. (2014). Brains creating stories of selves: The neural basis of autobiographical reasoning. *Social Cognitive and Affective Neuroscience, 9*, 646–652. http://dx.doi.org/10.1093/scan/nst028

Darwin, C. (1859/2002 reprinted). *On the origin of the species by means of natural selection, or the preservation of favoured races in the struggle for life. McLean, Virginia: IndyPublish.*

Davis, P. K., O'Mahony, A., Gulden, T. R., Osaba, O. A., & Sieck, K. (2018). *Priority challenges for social and behavioral research and its modeling.* Santa Monica, CA: RAND Corporation.

DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, *109*(8), E505–E514. http://doi.org/10.1073/pnas.1113427109

Dhami, M. K., & Belton, I. K. (2017). On getting inside the judge's mind. *Translational Issues in Psychological Science, 3*, 214–226. http://dx.doi.org/10.1037/tps0000115

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130*, 959–988. http://dx.doi.org/10.1037/0033-2909.130.6.959

Dhami, M. K., & Olsson, H. (2008). Evolution of the interpersonal conflict paradigm. *Judgment and Decision Making, 3*, 547–569.

Ekman, P. & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, California: Consulting Psychologists.

Eells, E. (1991). *Probabilistic causality*. Cambridge, England: Cambridge University Press.

Emmert, K., Kopel, R., Sulzer, J., Brühl, A. B., Berman, B. D., Linden, D. E. J., ... Haller, S. (2016). Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: How is brain regulation mediated? *NeuroImage, 124*, 806–812. http://dx.doi.org/10.1016/j.neuroimage.2015.09.042

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64. http://dx.doi.org/10.1518/001872095779049543

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

Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/9781316272503

Finkel, E. J., & Eastwick, P. W. (2008). Speed-dating. *Current Directions in Psychological Science, 17*, 193–197. http://dx.doi.org/10.1111/j.1467-8721.2008.00573.x

Fivush, R., Habermas, T., Waters, T. E.A., & Zaman, W. (2011). The making of autobiographical memory: Intersections of culture, narratives and identity. *International Journal of Psychology, 46*, 321–345. http://dx.doi.org/10.1080/00207594.2011.596541

Forgas, J. (1979). *Social episodes: The study of interaction routines.* London, England: Academic Press.

Fridman, M., & Petreanu, L. (2017), Cortical processing: How mice predict the visual effects. *Current Biology, 27* (23), 1268-1286.

Friston, K. J. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience 11*(2), 127–38.

Friston, K. J., & Kiebel, S. (2011). Predictive coding: A free-energy formulation. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 231–246). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195395518.003.0076

Friston, K. (2013). Active inference and free energy. *Behavioral and Brain Sciences,* 36, pp. 212-213. doi:10.1017/S0140525X12002142

Furl, N., Garrido, L, Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform Gurus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience, 23*(7), 1723-1740.

Gaba, D. M, Howard, S. K., Fish, K. J., Smith, B. E., & Sowb, Y. A. (2001). Simulation-based training in anesthesia crisis resource management (ACRM): A decade of experience. *Simulation Gaming. 32*, 175–193.

García-Betances, R. I., Arredondo Waldmeyer, M. T., Fico, G., & Cabrera-Umpiérrez, M. F. (2015). A succinct overview of virtual reality technology use in Alzheimer's disease. *Frontiers in Aging Neuroscience*, *7*, 80.

Garon, M., Boulet, P. O., Doironz, J. P., Beaulieu, L., & Lalonde, J. F. (2016, September). Real-time high resolution 3D data on the HoloLens. In 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct) (pp. 189-191). IEEE.

Garner, J. P., Gaskill, B. N., Weber, E. M., Ahloy-Dallaire, J., & Pritchett-Corning, K. R. (2017). Introducing Therioepistemology: the study of how knowledge is gained from animal research. *Lab Animal, 46*(4), 103–113. http://dx.doi.org/10.1038/laban.1224

Gaudiosi, J. (2016, January 5). Virtual reality video game industry to generate $5.1 billion in 2016. *Fortune*. Retrieved from http://fortune.com/2016/01/05/virtual-reality-game-industry-to-generate-billions/

Gee, J. P. (2003). *What video games have to teach us about learning and literacy.* New York: Palgrace Macmillan. ISBN 1-40396-169-7.

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science*, 14, 79-106.

Gendron, M., Mesquita, B., & Barrett, L. F. (2013). Emotion perception: Putting the face in context. In D. Reisberg (Ed.), *Oxford library of psychology: The Oxford handbook of cognitive psychology* (pp. 539–556). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780195376746.013.0034

Gergen, K. J. (1973). Social psychology as history. *Journal of Personality and Social Psychology, 26*, 309–320. http://dx.doi.org/10.1037/h0034436

Giacomin, M., & Jordan, C. H. (2014). Down-regulating narcissistic tendencies: Communal focus reduces state narcissism. *Personality and Social Psychology Bulletin, 40*(4), 488–500. doi:10.1177/0146167213516635.

Giacomin, M., & Jordan, C. H. (2016). The wax and wane of narcissism: Grandiose narcissism as a process or state. *Journal of Personality, 84*(2), 154–164. doi:10.1111/jopy.12148.

Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas (2008). *Psychological Review, 115*, 230–239. http://dx.doi.org/10.1037/0033-295X.115.1.230

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98*, 506–528. http://dx.doi.org/10.1037/0033-295X.98.4.506

Godoy, C. G., Miller, L. C., Corsbie-Massay, C., Christensen, J. L., Appleby, P. R., Read, S. J., & Si, Mei. (2013). Virtual Validity: mHealth Simulation Games, Diagnostic

Indicators, and Behavior Change. *Journal of Mobile Technology in Medicine, 2*(4S), 17. http://dx.doi.org/10.7309/jmtm.2.4s.14

Gore, W. L., & Widiger, T. A. (2016). Fluctuation between grandiose and vulnerable narcissism. *Personality Disorders: Theory, Research, and Treatment,7*(4), 363–371. doi:10.1037/per0000181.

Graesser, A. C. (1981). *Prose comprehension beyond the word*. New York, NY: Springer-Verlag.

Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American Psychologist*, *69*(1), 66-78. DOI: 10.1037/a0034857

Grant, H. & Dweck, C.S. (1999). A goal analysis of personality and personality coherence. In D. Cervone and Y. Shoda (Eds.) Social-cognitive approaches to personality coherence (pp. 345-371). New York: Guilford Press.

Grechuta, K., Ulysse, L., Ballester, B. R., & Verschure, P. F. (2019). Self beyond the body: action-driven and task-relevant purely distal cues modulate performance and body ownership. *Frontiers in Human Neuroscience*,*13*,91.

Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes, 38*, 247–266. http://dx.doi.org/10.1207/s15326950dp3802_5

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*, 701–721. http://dx.doi.org/10.1037/0022-3514.79.5.701

Hammond, K. R. (1965). New directions in research on conflict resolution. *Journal of Social Issues, 21*(3), 44–66. http://dx.doi.org/10.1111/j.1540-4560.1965.tb00505.x

Hammond, K. R., & Stewart, T. R. (2001). *The essential Brunswik: Beginnings, explications, applications*. New York, NY: Oxford University Press.

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis, *Military Psychology*, 4(2), 63-74.

Helmholtz, H. von (1860/1962) Handbuch der physiologischen optik, vol. 3, Ed. & trans. J. P. C. Southall. Dover. (Original work published in 1860; Dover English edition in 1962).

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2), 61–83. http://dx.doi.org/10.1017/s0140525x0999152x

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology, 15*(1), 1–39. http://dx.doi.org/10.1348/135910709x466063

Hieftje, K., Edelman, E. J., Camenga, D. R., & Fiellin, L. E. (2013). Electronic media–based health interventions promoting behavior change in youth: A systematic review. *JAMA Pediatrics*, *167*(6), 574-580. http://dx.doi.org/10.1001/jamapediatrics.2013.1095

Hilton, D. (2012). The emergence of cognitive social psychology: A historical analysis. In Kruglanski, A.W., and Stroebe, W. (Eds.) *Handbook of the history of social psychology.* New York, NY: Psychology Press (pages 45-80).

Hoffman, H. G., Richards, T. L., Coda, B., Bills, A. R., Blough, D., Richards, A. L., & Sharar, S. R. (2004). Modulation of thermal pain-related brain activity with virtual reality: Evidence from fMRI. *Neuroreport, 15*(8), 1245-1248.

Hoffman, H. G., Richards, T., Coda, B., Richards, A., & Sharar, S. R. (2003). The illusion

of presence in immersive virtual reality during an fMRI brain scan. *CyberPsychology & Behavior,* 6(2), 127-131.

Hoffman, P. M. (2012). *Life's ratchet: How molecular machines extract order from chaos.* New York: Basic Books. pp. 1-278.

Höll, M., Oberweger, M., Arth, C., & Lepetit, V. (2018, March). Efficient physics-based implementation for realistic hand-object interaction in virtual reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*(pp. 175-182). IEEE.

Hone-Blanchet, A., Wensing, T., & Fecteau, S. (2014). The use of virtual reality in craving assessment and cue-exposure therapy in substance use disorders. *Frontiers in Human Neuroscience, 8*, 844. http://dx.doi.org/10.3389/fnhum.2014.00844

Huang, C.-W., & Narayanan, S. S. (2017). Characterizing types of convolutional recurrent neural network with attention mechanism for robust speech emotion recognition. *2017 IEEE International Conference on Multimedia and Expo (ICME),* 1–18. http://dx.doi.org/10.1109/icme.2017.8019296

Hyatt, C. S., Sleep, C. E., Lynam, D. R., Widiger, T. A., Campbell, W. K., & Miller, J. D. (2018). Ratings of affective and interpersonal tendencies differ for grandiose and vulnerable narcissism: A replication and extension of Gore and Widiger (2016). *Journal of Personality*, *86*(3), 422-434.

Illari, P. M., Russo, F., & Williamson, J. (2011). Why look at causality in the sciences? A manifesto. In P. M. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 3–22). New York, NY: Oxford University Press. http://dx.doi.org/DOI:10.1093/acprof:oso/9780199574131.001.0001

Ito, R., & Lee, A. C. (2016). The role of the hippocampus in approach-avoidance conflict decision-making: Evidence from rodent and human studies. *Behavioural Brain Research*, *313*, 345-357.

Jeong, D. C., Feng, D., Krämer, N. C., Miller, L. C., & Marsella, S. (2017, August). Negative feedback in your face: examining the effects of proxemics and gender on learning. In *International conference on intelligent virtual agents* (pp. 170-183). Springer, Cham.

Johnsen, K., Raij, A., Stevens, A., Lind, D. S., & Lok, B. (2007, April). The validity of a virtual human experience for interpersonal skills education. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1049-1058). ACM.

Kelly, J., Gooding, P., Pratt, D., Ainsworth, J., Welford, M., & Tarrier, N. (2012). Intelligent real-time therapy: Harnessing the power of machine learning to optimise the delivery of momentary cognitive–behavioural interventions. *Journal of Mental Health, 21*(4), 404–414. http://dx.doi.org/10.3109/09638237.2011.638001

Kennedy, C. M., Powell, J., Payne, T. H., Ainsworth, J. H., Boyd, A., & Buchan, I. (2012). Active assistance technology for health-related behavior change: An interdisciplinary review. *Journal of Medical Internet Research, 14*(3), e80. http://dx.doi.org/10.2196/jmir.1893

Kenny, D. A., Mohr, C. D., & Levesque, M. J. (2001). A social relations variance partitioning of dyadic behavior. *Psychological Bulletin, 127*, 128-141.

Killingsworth, K., Miller, S. A., & Alavosius, M. P. (2016). A behavioral interpretation of situation awareness: Prospects for organizational behavior management. *Journal of Organizational Behavior Management, 36*(4), 301–321. http://dx.doi.org/10.1080/01608061.2016.1236056

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 2053951714528481.

Krebs, P., Prochaska, J. O., & Rossi, J. S. (2010). A meta-analysis of computer-tailored interventions for health behavior change. *Preventive Medicine, 51*(3-4), 214–221. http://dx.doi.org/10.1016/j.ypmed.2010.06.004

Krizan, Z., & Herlache, A. D. (2018). The narcissism spectrum model: A synthetic view of narcissistic personality. *Personality and Social Psychology Review, 22(*1), 3–31. doi:10.1177/1088868316685018.

Kveraga, K., Ghuman, A. & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition 65,* 145–68.

Lateef F. (2010). Simulation-based learning: Just like the real thing. *Journal of Emergencies, Trauma, and Shock*, *3*(4), 348–352. doi:10.4103/0974-2700.70743

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience, 8*, 302. http://dx.doi.org/10.3389/fnhum.2014.00302

Lee, K. M. (2004). Presence, explicated. *Communication Theory*, *14*(1), 27-50.

Lee, J., Sinclair, M., Gonzalez-Franco, M., Ofek, E., & Holz, C. (2019, April). TORC: A Virtual Reality Controller for In-Hand High-Dexterity Finger Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 71). ACM.

Levine, T. R., Blair, J. P., & Clare, D. D. (2013). Diagnostic utility: Experimental demonstrations and replications of powerful question effects in high-stakes deception detection. *Human Communication Research, 40*(2), 262–289. http://dx.doi.org/10.1111/hcre.12021

Li, B. J., Friston, K., Mody, M., Wang, H. N., Lu, H. B., & Hu, D. W. (2018). A brain network model for depression: From symptom understanding to disease intervention. *CNS Neuroscience & Therapeutics*. Advance online publication. http://dx.doi.org/10.1111/cns.12998

Lilienfeld, S. O. (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist, 67,* 111–129.

Lilienfeld, S. O. (2017). Psychology's Replication Crisis and the Grant Culture: Righting the Ship. *Perspectives on Psychological Science*, *12*(4), 660–664. https://doi.org/10.1177/1745691616687745

Limongi, R., Bohaterewicz, B., Nowicka, M., Plewka, A., & Friston, K. J. (2018). Knowing when to stop: Aberrant precision and evidence accumulation in schizophrenia. *Schizophrenia Research*, 197. Advance online publication. http://dx.doi.org/10.1016/j.schres.2017.12.018

Lindner, P., Miloff, A., Hamilton, W., Reuterskiöld, L., Andersson, G., Powers, M. B., & Carlbring, P. (2017). Creating state of the art, next-generation Virtual Reality exposure therapies for anxiety disorders using consumer hardware platforms: design considerations and future directions. *Cognitive Behaviour Therapy*, *46*(5), 404-420.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-mediated Communication*, *3*(2), JCMC 321.

Lombard, M., & Jones, M. T. (2015). Defining presence. In *Immersed in Media*(pp. 13-34). Springer, Cham.

Mandler, J. M. (1978). A code in the node: The use of story schema in retrieval. *Discourse Processes, 1*, 14–35. Retrieved from http://doi.org/10.1080/01638537809544426

Mann, R. (2010). An introduction to particle physics and the Standard Model. Baco Raton, Florida: CRC Press. ISBN 978-1-4200-8298-2.

Maples-Keller, J. L., Bunnell, B. E., Kim, S. J., & Rothbaum, B. O. (2017). The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard Review of Psychiatry*, *25*(3), 103.

Marsella, S., & Gratch, J. (2016). Computational models of emotion as psychological tools. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (Fourth ed., pp. 113–132). New York, NY: Guilford Press.

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

Marsella, S. C., Johnson, W. L., & LaBore, C. (2000). Interactive pedagogical drama. In C. Sierra, M. Gini, & J. S. Rosenschein (Eds.), *Proceedings of the Fourth International Conference on Autonomous Agents* (pp. 301–308). New York, NY: ACM. http://dx.doi.org/10.1145/336595.337507

Marsella, S. C., Pynadath, D. V., & Read, S. J. (2004). PsychSim: Agent-based modeling of social interactions and influence. In M. Lovett, C. Schunn, C. Lebiere, & P. Munro (Eds.), *Proceedings of the Sixth International Conference on Cognitive*

Modeling: ICCCM 2004: Integrating Models (pp. 243–248). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

McAdams, D. P. (1990). Unity and purpose in human lives: The emergence of identity as a life story. In A. I. Rabin, R. A. Zucker, R. A. Emmons, & S. Frank (Eds.), *Studying persons and lives* (pp. 148–200). New York, NY: Springer Publishing Co.

McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology, 5*(2), 100–122. http://dx.doi.org/10.1037//1089-2680.5.2.100

McAdams, D. P. (2008). Personal narratives and the life story. In O. John, R. Robins, and L. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd Ed.) (pp. 241-261). New York: Guilford Press.

McClelland, J. L., & Elman J. L. (1986). Interactive processes in speech perception. The TRACE model. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models* (pp. 58-121). Cambridge MA: MIT Presss/Bradford Books.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375-407.

Mesulam MM (1998). From sensation to cognition. *Brain: A Journal of Neurology, 121*, 1013–1052.

Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations, 18(1)*, 57-76.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY: Henry Holt and Co. http://dx.doi.org/10.1037/10039-000

Miller, J. D., Lynam, D. R., Hyatt, C. S., & Campbell, W. K. (2017). Controversies in narcissism. *Annual Review of Clinical Psychology*, 13(1), 291-315. https://doi.org/10.1146/annurev-clinpsy-032816-045244

Miller, L. C., Christensen, J. L., Godoy, C. G., Appleby, P. R., Corsbie-Massay, C., & Read, S. J. (2009). Reducing risky sexual decision-making in the virtual and in the real-world: Serious games, intelligent agents, and a SOLVE approach. In U. Ritterfeld, M. Cody, P. Vorderer (Eds.) *Serious games: Mechanisms and effects.* (pp 34–37). New York, NY: Routledge/LEA Press.

Miller, L.C., Cody, M.J., & McLaughlin, M.L. (1994). Goals and situations as fundamental constructs in interpersonal communication research. In M.L. Knapp and G.R. Miller (Eds.), *Handbook of interpersonal communication* (pp. 162-198). Newbury Park, California: Sage.

Miller, L. C., Jeong, D. C., Christensen, J. L. (2019). The Promise of Virtual Intelligent Games for Assessing Personality. In Wood, D., Read, S. J., Harms, P. D., & Slaughter, A. (Eds.). *Measuring and modeling persons and situations;* manuscript in preparation, will be published under Elsevier's *Academic Press* imprint.

Miller, L. C., Marsella, S., Dey, T., Appleby, P. R., Christensen, J. L., Klatt, Jennifer, & Read, Stephen J. (2011). Socially optimized learning in virtual environments (SOLVE). In M. Si, D. Thue, E. Andre, J. Lester, J. Tannenbaum, & V. Zammitto (Eds.), *Interactive Storytelling: Proceedings of the Fourth International Conference on Interactive Digital Storytelling* (pp. 182–192). New York, NY: Springer.

Miller, L. C., & Read, S. J. (1987). Why am I telling you this?: Self-disclosure in a goal-based model of personality. In V. J. Derlega & J. H. Berg (Eds.), *Perspectives in social psychology: Self-disclosure theory, research, and therapy* (pp. 35–58). New York, NY: Plenum Press.

Miller, L. C., & Read, S. J. (1991). Inter-personalism: Understanding persons in relationships. In W. Jones & D. Perlman (Eds.), *Advances in personal relationships* (Vol. 2, pp. 233–267). London, England: Kingsley.

Miller, L. C., Wang, L., Jeong, D., & Gillig, T. (2019). Bringing the "real-world" into the experimental lab: Technology-enabling transformative designs. In P. Davis, A. O'Mahony, & J. Pfautz (Eds.), *Social behavioral science modeling*. Hoboken, NJ.: Wiley-Blackwell.

Miller, N. E. (1944). Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 431--465). New York: Ronald Press.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38*(4), 379–387. http://dx.doi.org/10.1037//0003-066x.38.4.379

Morf, C. C. (2006). Personality reflected in a coherent idiosyncratic interplay of intra- and interpersonal self-regulatory processes. *Journal of Personality, 74*(6), 1527–1556. doi:10.1111/j.1467-6494.2006.00419.x.

Morris, Z. S., Wooding, S., & Grant, J. (2011). The answer is 17 years, what is the question: Understanding time lags in translational research. *Journal of the Royal Society of Medicine, 104*(12), 510–520.

Moser, E. I., Kropff, E., & Moser, M-B (2008). Place cells, grid cells, and the brain's spatial representation system, *Annual Review of Neuroscience, 31,* 69-89. https://doi.org/10.1146/annurev.neuro.31.061307.090723

Mueller, C., Luehrs, M., Baecke, S., Adolf, D., Luetzkendorf, R., Luchtmann, M., & Bernarding, J. (2012). Building virtual reality fMRI paradigms: a framework for presenting immersive virtual environments. *Journal of Neuroscience Methods*, *209*(2), 290-298.

Myerson, Roger B. (1991). *Game Theory: Analysis of conflict,* Boston, Mass.: Harvard University Press.

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology, 34*(Suppl), 1209–1219. http://dx.doi.org/10.1037/hea0000306

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine, 52*(6), 446–462. http://dx.doi.org/10.1007/s12160-016-9830-8

National Academy of Sciences, Engineering, and Medicine (2019); Reproducibility and replicability in science. Washington, DC: *The National Academies Press*. https://doi.org/10.17226/25303.

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review, 111*(2), 486–511. http://dx.doi.org/10.1037/0033-295X.111.2.486

Nobel Media AB (1998, October 13). *The Nobel Prize in chemistry 1998*.Retrieved May 20, 2019 from https://www.nobelprize.org/prizes/chemistry/1998/summary/

Nobel Media AB (1999, October 13). *The Nobel Prize in physics 1999*. Retrieved May 20, 2019 from https://www.nobelprize.org/prizes/physics/1999/summary/

Nobel Media AB (2010, October 5). *The Nobel Prize in physics 2010*. Retrieved May 20, 2019 from https://www.nobelprize.org/prizes/physics/2010/summary/

Nobel Media AB (2012, October 9). *The Nobel Prize in physics 2012*. Retrieved May 20, 2019 from https://www.nobelprize.org/prizes/physics/2012/prize-announcement/

Nobel Media AB (2013, October 9). *The Nobel Prize in chemistry 2013*.Retrieved May 20, 2019 from https://www.nobelprize.org/prizes/chemistry/2013/summary/

Nobel Media AB (2014, October 8). *The Nobel Prize in chemistry 2014*.Retrieved May 20,2019 fromhttps://www.nobelprize.org/prizes/chemistry/2014/summary/

Nobel Media AB (2018a). *Frances H. Arnold – Facts – 2018*. Retrieved May 20, 2019 from https://www.nobelprize.org/prizes/chemistry/2018/arnold/facts

Nobel Media AB (2018b, October 2). *Press release: The Nobel Prize in physics 2018*. Retrieved May 20, 2019 from https://www.nobelprize.org/prizes/physics/2018/press-release/

Oehrn, C. R., Baumann, C., Fell, J., Lee, H., Kessler, H., Habel, U., Hanslmayr, S., & Axmacher, N. (2015). Human hippocampal dynamics during response conflict. *Current Biology*, *25*(17), 2307-2313.

O’Keefe, J., & Nadel, L. (1978). The hippocampus as a cognitive map. Oxford, England: Oxford University Press.

O'Neil, E. B., Newsome, R. N., Li, I. H., Thavabalasingam, S., Ito, R., & Lee, A. C. (2015). Examining the role of the human hippocampus in approach–avoidance decision making using a novel conflict paradigm and multivariate functional magnetic resonance imaging. *Journal of Neuroscience*, *35*(45), 15039-15049.

Overmars, M. (2005). Learning object-oriented design by creating games. *IEEE Potentials, 23*, 11–13. http://dx.doi.org/10.1109/mp.2005.1368910

Owens, A. P., Allen, M., Ondobaka, S., & Friston, K. J. (2018). Interoceptive inference: From computational neuroscience to clinic. *Neuroscience and Biobehavioral Reviews, 90*, 174–183. http://dx.doi.org/10.1016/j.neubiorev.2018.04.017

Paluck, E. L., & Cialdini, R. B. (2014). Field research methods. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 81–97). New York, NY: Cambridge University Press.

Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: A literature review. *Computers & Education, 53*(3), 603–622. http://dx.doi.org/10.1016/j.compedu.2009.04.001

Parsons, T., Gaggioli, A., & Riva, G. (2017). Virtual reality for research in social neuroscience. *Brain Sciences*, *7*(4), 42.

Pearl, J., & Bareinboim, E. (2014). External validity: From do-calculus to transportability across populations. *Statistical Science, 29*(4), 579–595. http://dx.doi.org/10.1214/14-STS486

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. New York, NY: Basic Books.

Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology, 51*, 242–258. http://dx.doi.org/10.1037/0022-3514.51.2.242

Pettigrew, T. F. (2018). The emergence of contextual social psychology. *Personality and Social Psychology Bulletin, 44*, 963–971. http://dx.doi.org/10.1177/0146167218756033.

Pincus, A. L. , Cain, NM, & Wright, A. G. (2014). Narcissistic grandiosity and narcissistic vulnerability in psychotherapy. *Personality Disorders: Theory, Research, and Treatment,* 5(4), 439-443. http://dx.doi.org/10.1037/per0000031.

Pincus, A. L., Roche, M. J., & Good, E. W. (2015). Narcissistic personality disorder and pathological narcissism. In P. H. Blaney, R. F. Krueger, and T. Millon (Eds.), Oxford textbook of psychopathology (3rd ed. pp. 791–813). New York: Oxford University Press.

Price, M., & Anderson, P. (2007). The role of presence in virtual reality exposure therapy. *Journal of Anxiety Disorders*, *21*(5), 742-751.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59.

Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(6), 608–620. http://doi.org/10.1002/wcs.133

Radvansky, G. A., & Zacks, J. M. (2014). *Event cognition*. Oxford ; New York: Oxford University Press.

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences*, *97*(22), 11800–11806. http://doi.org/10.1073/pnas.97.22.11800

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., ... Funder, D. C. (2014). The situational eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology, 107*(4), 677–718. http://dx.doi.org/doi.org/10.1037/a0037250

Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. *Journal of Personality and Social Psychology, 52*, 288. http://dx.doi.org/10.1037/0022-3514.52.2.288

Read, S. J., Droutman, V., Smith, B. J., & Miller, L. C. (2017). Using neural networks as models of personality process: A tutorial. *Personality and Individual Differences*. 136(1), 52-67. http://dx.doi.org/10.1016/j.paid.2017.11.015

Read, S. J., Druian, P. R., & Miller, L. C. (1989). The role of causal sequence in the meaning of actions. *British Journal of Social Psychology, 28*, 341–351. http://dx.doi.org/10.1111/j.2044-8309.1989.tb00877.x

Read, S. J., & Miller, L.C. (1989a). The importance of goals in personality: Toward a coherent model of persons. In R.S. Wyer, Jr., & T.K. Srull (Eds.), Advances in social cognition, Volume 2: *Social intelligence and cognitive assessments of personality.* Hillsdale, N.J.: Erlbaum.

Read, S.J., & Miller, L. C. (1989b). Inter-personalism: Toward a goal-based model of persons in relationships. In L. Pervin (Ed.), *Goal concepts in personality and social psychology.* Hillsdale, N.J.: Erlbaum.

Read, S. J., & Miller, L. C. (1995). Stories are fundamental to meaning and memory: For social creatures could it be otherwise? In R. S. Wyer, Jr. (Ed.), *Advances in social cognition: Vol. 8. Knowledge and memory: The real story* (pp. 139–152). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Read, S. J., & Miller, L. C. (Eds.). (1998). *Connectionist models of social reasoning and social behavior*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Read, S. J., & Miller, L. C. (in press). A neural network model of motivated decision-making and everyday behavior. In P. Davis, A. O'Mahony, & J. Pfautz (Eds.), *Social behavioral science modeling*. Hoboken, NJ: Wiley.

Read, S. J., Miller, L. C., Appleby, P. R., Nwosu, M. E., Reynaldo, S., Lauren, A., & Putcha, A. (2006). Socially Optimized Learning in a Virtual Environment: Reducing Risky Sexual Behavior Among Men Who Have Sex With Men. *Human Communication Research, 32*, 1–34. http://dx.doi.org/10.1111/j.1468-2958.2006.00001.x

Read, S. J., Monroe, B. M., Brownstein, A. L., Yang, Y., Chopra, G., & Miller, L. C. (2010). A neural network model of the structure and dynamics of human personality. *Psychological Review, 117*, 61–92. http://dx.doi.org/10.1037/a0018131

Read, S. J., Smith, B. J., Droutman, V., & Miller, L. C. (2017). Virtual personalities: Using computational modeling to understand within-person variability. *Journal of Research in Personality, 69*, 237–249. http://dx.doi.org/10.1016/j.jrp.2016.10.005

Read, S. J., Brown, A. B., Wang, P., & Miller, L. C. (2018). The Virtual Personalities Neural Network Model: Neural Biological Underpinnings. *Personality Neuroscience. Vol 1*: e10, 1–11. doi:10.1017/ pen.2018.6

Read, S. J., Wang, P., Brown, A. D., Smith, B. J., & Miller, L. C. (in press). Neural networks and virtual personalities. In J. Rathmann (Ed.), *The handbook of personality dynamics and processes*. Amsterdam, the Netherlands: Elsevier.

Regenbrecht, H. T., Schubert, T. W., & Friedmann, F. (1998). Measuring the sense of presence and its relations to fear of heights in virtual environments. *International Journal of Human-Computer Interaction*,*10*(3), 233-249.

Reis, H. T. (2008). Reinvigorating the concept of situation in social psychology, *Personality and Social Psychology Review, 12*(4), 311-329.

Reese, E. (2002). A model of the origins of autobiographical memory. In J. W. Fagen & H. Hayne (Eds.), *Progress in infancy research* (Vol. 2, pp. 215–260). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Rescorla, R.A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.) *Classical conditioning II: Current research and theory* (pp. 64-98). New York: Appleton-Century-Crofts.

Ritterfeld, U., Cody, M., & Vorderer, P. (Eds.). (2009). *Serious games: Mechanisms and effects*. New York, NY: Routledge.

Riva, G., Wiederhold, B. K., & Mantovani, F. (2019). Neuroscience of virtual reality: From virtual exposure to embodied medicine. *Cyberpsychology, Behavior, and Social Networking, 22* (1), 82-96. DOI: 10.1089/cyber.2017.29099.gri

Rizzo, A. S., & Koenig, S. T. (2017). Is clinical virtual reality ready for primetime? *Neuropsychology, 31*, 877–899. http://dx.doi.org/10.1037/neu000040

Rizzo, A. S., Hartholt, A., Rothbaum, B., Difede, J., Reist, C., Kwok, D., ... Buckwalter, J. G. (2014). Expansion of a VR exposure therapy system for combat-related PTSD to Medics/Corpsman and persons following military sexual trauma. *Studies in Health Technology and Informatics, 196*, 332–338. http://dx.doi.org/10.3233/978-1-61499-375-9-332

Robineau, F., Meskaldji, D. E., Koush, Y., Rieger, S. W., Mermoud, C., Morgenthaler, S., ... Scharnowski, F. (2017). Maintenance of voluntary self-regulation learned through real-time fMRI neurofeedback. *Frontiers in Human Neuroscience, 11*, 131. http://dx.doi.org/10.3389/fnhum.2017.00131

Roseman, I. J. (2008). Motivations and Emotivations: Approach, avoidance, and other tendencies in motivated and emotional behavior. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation*(pp. 343–366). New York NY: Psychology Press.

Roseman, I. J. (2011). Emotional behaviors, emotional goals, emotion strategies: Multiple levels of organization integrate variable and consistent responses. *Emotion Review, 3*(4), 434-443. DOI:10.1177/1754073911410744 er.sagepub.com

Rothbaum, B. O., Rizzo, A., & Difede, J. (2010). Virtual reality exposure therapy for combat-related posttraumatic stress disorder. *Annals of the New York Academy of Sciences*, *1208*(1), 126-132.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. LaBerge & J. Samuels (Eds.), *Basic processes in reading and comprehension* (pp. 265–303). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests of extensions of the model. *Psychological Review, 89*, 60-94.

Scerri, E. R. (2007). *The periodic table: Its story and its significance.* New York, Oxford University Press.

Schank, R. C., & Abelson, R. P. (1977). *The Artificial Intelligence Series: Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Schank, R. C., & Abelson, R. P. (1995). Knowledge and memory: The real story. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition: Vol. 8. Knowledge and memory: The real story* (pp. 1–85). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006).Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718-730.

Schönbrodt, F. D., & Asendorpf, J. B. (2011). Virtual social environments as a tool for psychological assessment: Dynamics of interaction with a virtual spouse. *Psychological Assessment*, *23*(1), 7-17.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51*(3), 515–530. http://dx.doi.org/10.1037//0022-3514.51.3.515

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton, Mifflin and Company.

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments*,*1*(1), 120-126.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1-32. https://doi.org/10.1146/annrec.clinpsy.3022806.091415

Si, M., Marsella, S., & Pynadath, D. (2010). Importance of well-motivated characters in interactive narratives: an empirical evaluation. In R. Aylett (Ed.), *Joint International Conference on Interactive Digital Storytelling* (pp. 16–25). Berlin, Germany: Springer-Verlag.

Smith, B. J., Xue, F., Droutman, V., Barkley-Levenson, E., Melrose, A. J., Miller, L. C., ... Read, S. J. (2018). Virtually 'in the heat of the moment': Insula activation in safe sex negotiation among risky men. *Social Cognitive and Affective Neuroscience, 13*, 80–91. http://dx.doi.org/10.1093/scan/nsx137

Somandepalli, K., Gupta, R., Nasir, M., Booth, B. M., Lee, S., & Narayanan, S. S. (2016, October). Online affect tracking with multimodal kalman filters. In *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge*(pp. 59-66). ACM.

Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and Cognition*, 112, 92-97. https://doi.org/10.1016/j.bandc.2015.11.003

Spruijt-Metz, D., Wen, C. K. F., O'Reilly, G., Li, M., Lee, S., Emken, B. A., ... Narayanan, S. (2015). Innovations in the Use of Interactive Technology to

Support Weight Management. *Current Obesity Reports, 4*, 510–519. http://dx.doi.org/10.1007/s13679-015-0183-6

Steinicke, W. (2005). Einstein and the gravitational waves. *Astronomische Nachrichten, AN,* 326 (7), 1.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, *42*(4), 73-93.

Swann, W. B., Jr., & Seyle, C. (2005). Personality psychology's come- back and its emerging symbiosis with social psychology. *Personality and Social Psychology Bulletin, 31,* 155–165.

Tepper, O. M., Rudy, H. L., Lefkowitz, A., Weimer, K. A., Marks, S. M., Stern, C. S., & Garfein, E. S. (2017). Mixed reality with HoloLens: where virtual reality meets augmented reality in the operating room. *Plastic and reconstructive surgery*, 140(5), 1066-1070.

Thompson, D. (2012). Designing serious video games for health behavior change: Current status and future directions. *Journal of Diabetes Science and Technology*, 6(4), 807-811. doi: 10.1177/193229681200600411

Thorne, A., & McLean, K. C. (2003). Telling traumatic events in adolescence: A study of master narrative positioning. In R. Fivush & C. A. Haden (Eds.), *Autobiographical memory and the construction of a narrative self: Developmental and cultural perspectives* (pp. 169–185). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N., & Falvello, V. B. (2013). Validation of data-driven computational models of social perception of faces. *Emotion, 13*(4), 724-738. doi:http://dx.doi.org.libproxy1.usc.edu/10.1037/a0032335

Tokareva, J. (2018, February 02). The Difference Between Virtual Reality, Augmented Reality And Mixed Reality. *Forbes*. Retrieved May 9, 2019, from https://www.forbes.com/sites/quora/2018/02/02/the-difference-between-virtual-reality-augmented-reality-and-mixed-reality/#115378622d07

Traylor, A. C., Parrish, D. E., Copp, H. L., & Bordnick, P. S. (2011). Using virtual reality to investigate complex and contextual cue reactivity in nicotine dependent problem drinkers. *Addictive Behaviors, 36*(11), 1068–1075. http://dx.doi.org/10.1016/j.addbeh.2011.06.014

Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013). Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics? *Current Directions in Psychological Science*, *22*(5), 367-373.

Vallacher, R. R., Read, S. J., & Nowak, A. (Eds.). (2017). *Frontiers of social psychology: Computational social psychology*. New York, NY: Routledge.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences, 113*, 6454–6459. http://dx.doi.org/10.1073/pnas.1521897113 van Bennekom, M. J., Kasanmoentalib, M. S., de Koning, P. P., & Denys, D. (2017). A virtual reality game to assess obsessive-compulsive disorder. *Cyberpsychology, Behavior, and Social Networking*, *20*(11), 718-722.

Van Den Bos, E., & Jeannerod, M. (2002). Sense of body and sense of action both contribute to self-recognition. *Cognition*, *85*(2), 177-187.

Van Overwalle, F. (1997). A test of the joint model of causal attribution. *European Journal of Social Psychology, 27*, 221-236.

Van Overwalle, F. & Van Rooy, D. (1998). A connectionist approach to causal attribution. In S. Read & L.Miller (Eds.). *Connectionist models of social reasoning and social behavior* (pp. 143-174). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Wang, L., & Miller, L. C. (in press). Do just in-the-moment, adaptive interventions (JITAI) work? A Meta-Analysis. *Health Communication*.

Weibel, R. P., Grübel, J., Zhao, H., Thrash, T., Meloni, D., Hölscher, C., Schinazi, V. R. (2018). Virtual reality experiments with physiological measures. *Journal of Visualized Experiments*, *138*, e58318, doi:10.3791/58318.

Wiederhold, B. K., & Wiederhold, M. D. (2008). Virtual reality with fMRI: A breakthrough cognitive treatment tool. *Virtual Reality, 12,* 259-267. DOI: 10.1007/s10055-008-0100-3.

Wierzbicka, A. (1992). *Semantics, culture, and cognition: Universal human concepts in culture-specific configurations*. Oxford University Press on Demand.

Wilson, T. D. (2012, July 12). Stop bullying the ‘soft’ sciences. *LA Times*. Retrieved May 20, 2019 from https://www.latimes.com/opinion/la-xpm-2012-jul-12-la-oe-wilson-social-sciences-20120712-story.html

Witkower, Z., & Tracy, J. L. (2019). A facial-action imposter: How head tilt influences perceptions of dominance from a neutral face. *Psychological Science*, *30*(6), 893-906. https://doi.org/10.1177/0956797619838762.

Wittgenstein, L., 2009. *Philosophical Investigations.* Translated by Anscombe, G.E.M., Hacker, P.M.S., and Schulte, J. from Philosophische Untersuchen (1953). Hoboken: Wiley-Blackwell.

Wundt, W. (1902). *Outlines of psychology* (2nd ed.). Oxford, England: Engelmann.

Yang, Y., Read, S. J., & Miller, L. C. (2006). A taxonomy of situations from Chinese idioms. *Journal of Research in Personality, 40*, 750–778. http://dx.doi.org/10.1016/j.jrp.2005.09.007

Zook, M., Barocas, S., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B.A., Metcalf, J., Narayanan, A., Nelson, A., & Pasquale, F. (2017). Ten simple rules for responsible big data research. *PLoS Computational Biology* 13(3): e1005399.

Table 1

Comparing Three Designs on Inference Goals and Strategies to Achieve Them

***Goals* & Strategies**	**Systematic Design**	**Representative Design**	**Systematic Representative Design**
Cause-Effect			
Study Phase	Design	Causal Analysis	Design + Analysis
Timing	IV -> DV	Sometimes	IV -> DV
Controls	Via Manipulation	Via Analysis	Via Manipulation
Eliminate Alternative Explanations	Random Assignment	Analysis	Random Assignment
	Control (CG) and Experimental (EG)	Analysis	Default Control (DCG) + EG on DCG
Generalizability			
Everyday Life (GEL); Ext. Valid (EV)	EV, but Even If EV, GEL Unknown	GEL; If GEL, Then EV Likely	GEL; If GEL, Then EV Likely
Representative?	Atypical	Task Analysis	Task Analysis
Check?	No	No	Virtual Validity
Denatured Variable?	Denatured Variables Tying/Untying	Naturally Occurring	Naturally Occurring

Table 2

Representatively Designing a Default Control Group: Social Scientist Steps for Intelligent Agent/Game Designer Collaborator in Building Default Control Group

Step	Concept	Example
1	Identify Behavior(s) of Interest (BOI)	Condomless Anal Sex (CAS)
2	Identify Samples of Population of Interest (POI)	Young Men who have Sex with Men (YMSM)
3	Identify Most Frequent Settings (MFS) Leading Up to BOI for POI	Home/Apartment (CAS frequently occurs); House Party; Bar/Club; Internet (First Contact)
4	Extract Details, Relevant Cues	Cues for POI in MFS; Partner Selection Attributes
5	Identify Scripts in MFS	"Pick Up"; "Sexual Script"
6	Identify Components of Scripts (SC)	Bar Pick-Up Steps: Enter/Check Scene; Zero in on Prospective Target; Create Reason to Meet; Get to Know; Test waters; Escalate Intimacy; Seal the deal
7	"Entry" and "Exit" Conditions for Each SC and Specified (Extract Relevant Cues to Threshold Exit/Entry Conditions. Identify Frequent Responses/ Options; How Sequences can go Differently	Bar Pick-Up from "Test Waters" to "Escalate Intimacy" Example: (1) Pretext to Touch (e.g., Love the Feel of your Shirt, is it Silk?); If go, (2) Reduce Distance (e.g., Dance) (3) Brush up "Accidently" (Deniability) Until more Intimate (e.g., Touch Leg); If Go, Exit (4) Enter Escalate Intimacy (more Direct,

		Foreplay), Desire for Sex Mutual, if Threshold Exit; (5) Enter "Seal the Deal").
8	Identify Challenges (e.g., to Safer Sex) Embedded in Sequence up to BOI for POI; When/How they Occur	Frequent Obstacles (e.g., Alcohol) Given BOI and POI. In "Getting to Know" Phase, Offered Drink.
9	Represent Exact Behavioral Implementations for Dialogue for Agent; Representative Human Response Options Where Variability	In "Getting to Know," Compliments Afford Positivity Threshold. Specify for POI, Positively Rated Compliments in Bar Scenario. Specify Behavior Options; Choice Basis (e.g., in Attachment Styles)

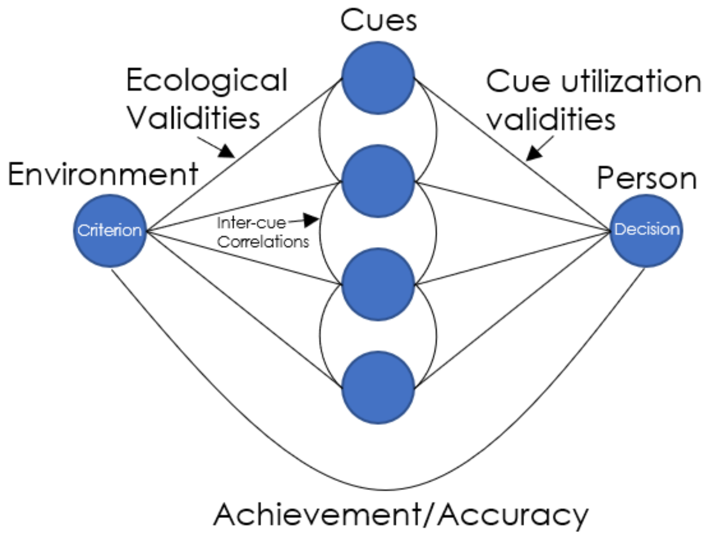

Figure 1. Adapted from Brunswik's Lens Model of a Single-System (1952).

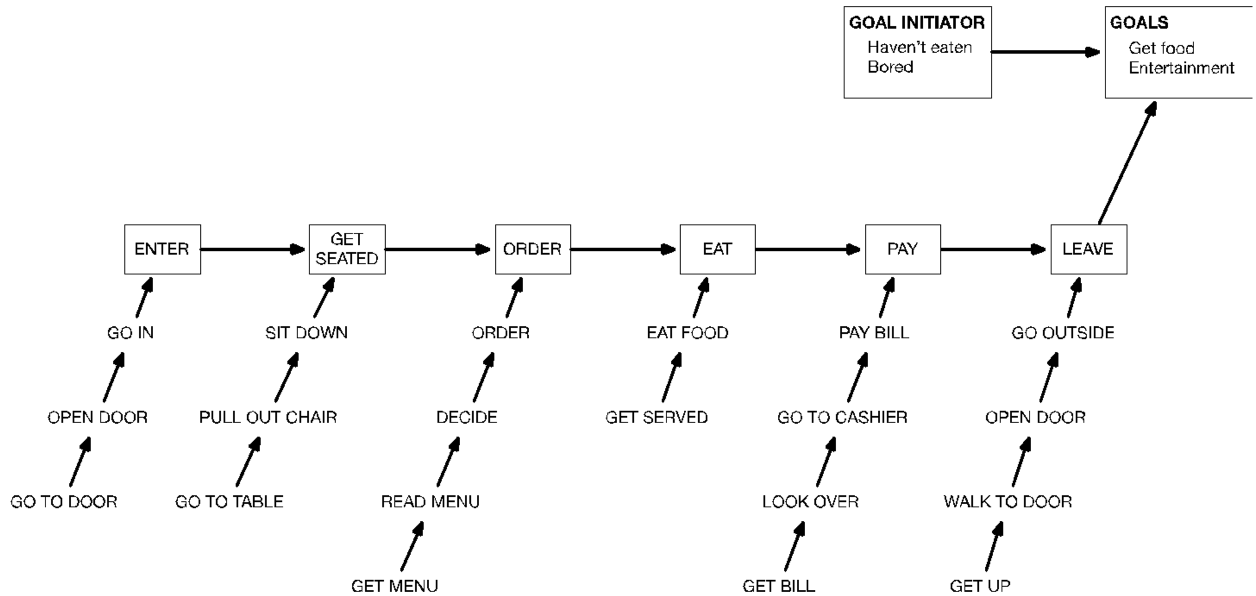

Figure 2. Restaurant Script Adapted from Read (1987).

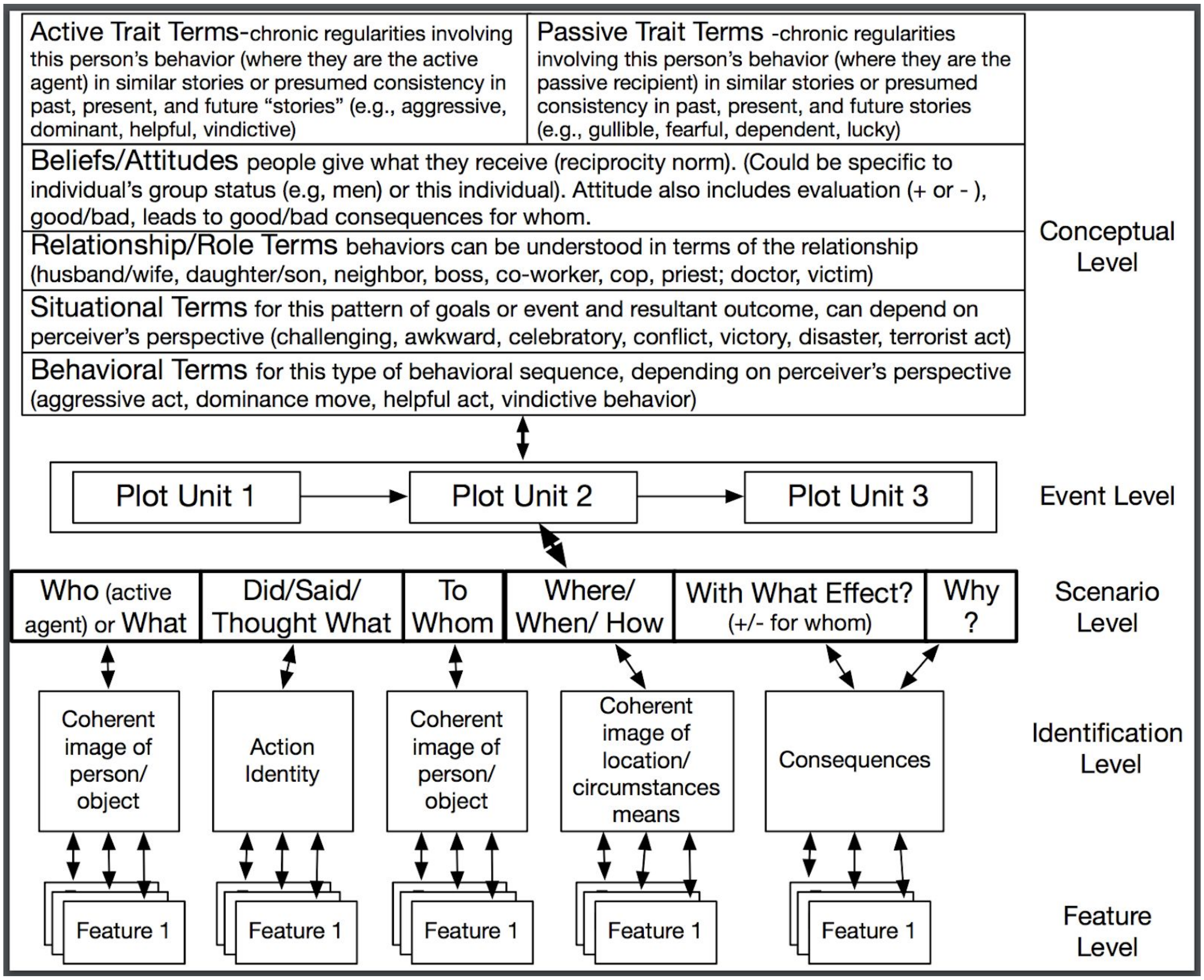

Figure 3.

The bi-directional information flow between the levels of the Constraint Satisfaction Model exemplified (modified from Read & Miller, 1998 by adding the event level for clarification).

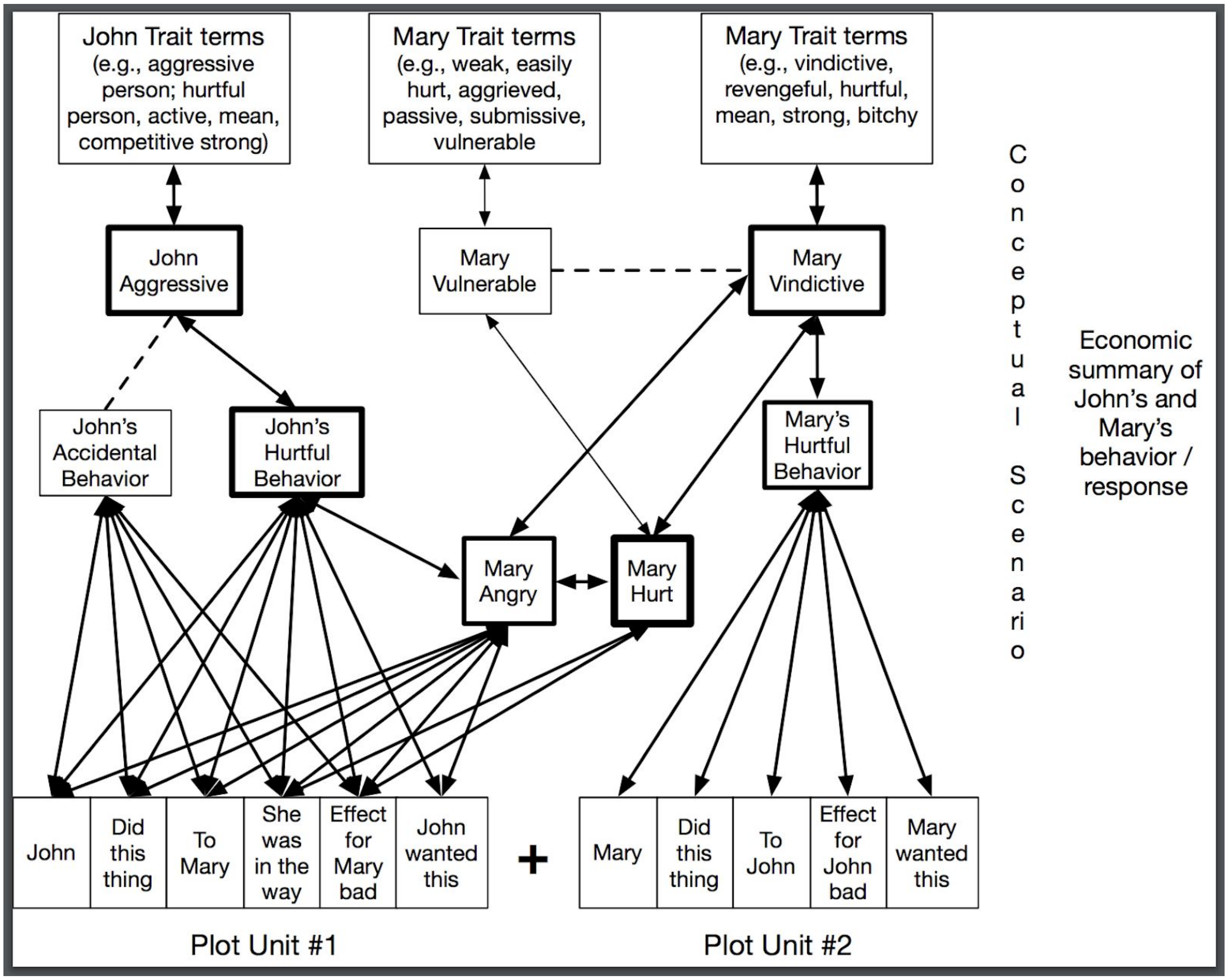

Figure 4. Constraint Satisfaction Model of Social Perception, Specific Example