Simulation induces durable, extensive changes to self-knowledge☆

Jordan Rubin-McGregor [*], Zidong Zhao, Diana I. Tamir

Department of Psychology, Princeton University, Princeton, NJ 08540, United States of America

ARTICLE INFO

ABSTRACT

The sense of self is a hallmark of the human experience, but it is also unstable. Even simulating another person – thinking about their traits or experiences – can shift how one thinks about their own traits or experiences. Simulating a target shifts self-knowledge such that it becomes more similar to the target; in six studies, we explore how extensively these changes occur. In all studies, participants first rated themselves in a specific context, then simulated another individual in the same context, and finally considered themselves again. We calculated how participants' self-knowledge changed by comparing similarity to the target before vs. after simulation. In Studies 1–2, participants' episodic memories shifted to be more similar to the simulated target; this change persisted at least 48 h. Studies 3–4 show that semantic self-knowledge changes after considering semantically related traits, while Study 5 shows that this effect extends to cross-language traits. Together, these results suggest that simulation causes durable, extensive changes across both episodic and semantic self-knowledge.

1. Introduction

What makes us 'us'? This question has animated psychologists for centuries. In the early 20th century, Calkins (1908) had already defined psychology as the "science of self." As early as 1902, Cooley proposed that the self is shaped by its social surroundings. Since then, psychologists have indeed shown that the self is not a static entity, waiting to be discovered within each of us. Instead, one's self can change considerably over development, across situations, after major life events, or after minimal social feedback (Downey, Rosengren, & Donovan, 2000; Gore, 2005; Markus & Wurf, 1987). The self is a key construct for any individual and for the whole of psychological science to understand, and yet, how it takes shape over time is not fully understood. We know that the self can change, but here, we ask: How much? And through what forces?

1.1. Forces that change the self

The sense of self has core components that tend to remain stable over time (Demo, 1992; Diehl, Jacobs, & Hastings, 2006; Trzesniewski, Donnellan, & Robins, 2003). Despite this backdrop of stability, however, self-knowledge can change in a myriad of ways. For example, one's sense of self changes over time, from childhood to adulthood (Bloom, 1961; McCrae & Costa, 1988). Multiple aspects of the self can be changed,

including episodic memories, which are personally experienced past events (Tulving, 2002), and semantic self-knowledge, or the content that defines what individuals think and believe about themselves (Purkey, 1988), such as one's perceptions of their own characteristics (Sakaki, 2007b). While there are many additional elements that comprise the self, such as self-esteem or group identification, here we focus on the forces that change autobiographical episodic memories and semantic self-knowledge, in particular (Ellemers, Kortekaas, & Ouwerkerk, 1999), as two key components of many models of self-concept (Conway, 2005; DeSteno & Salovey, 1997; Sakaki, 2007a).

Multiple factors can shift self-concept. External factors, particularly in the form of direct social interaction, are a major catalyst for change (Cooley, 1902). For example, after being described as dominant or submissive by a conversation partner, people who are not given a chance to deny the assessment will shift their self-concept to match that feedback (Swann & Hill, 1982). When individuals are asked leading questions that imply introversion or extroversion, self-ratings on those traits will shift to match the leading questions (Fazio, Effrein, & Falender, 1981). Direct social interaction offers people new information about themselves that then shifts their self-concept. Self-concept can also shift as a result of the situations individuals are in. For example, people's self-ratings on the Big Five traits change depending on their environment. This change can be reliably predicted based on characteristics of the

situation (i.e., the number of people present, how interesting the situation is, etc.; Fleeson, 2007); the frequency of positive daily experiences can likewise predict changes in the Big Five traits, with more positive experiences resulting in higher ratings of extraversion, conscientiousness, emotional stability, and openness (Borghuis et al., 2018). Traits also change across contexts, above and beyond what can be accounted for by changes in affect; for example, individuals show higher self-ratings on conscientiousness when studying or working than when in other situations (Wilson, Thompson, & Vazire, 2017). Thus, even without direct feedback from social interactions, external factors can elicit changes in self-concept.

In addition to external sources, purely internal forces can cause changes to the self. That is, people don't need the external environment to provide feedback or new information for the self to change; people can generate that information themselves. Imagination, though internal, is a powerful force for shaping behavior and cognition. When individuals vividly imagine climate change, their pro-environmental goals and behaviors increase (e.g., recycling; Boomsma, Pahl, & Andrade, 2016). When individuals vividly imagine reaching a goal, they increase goal-directed behavior (Taylor, Pham, Rivkin, & Armor, 1998). Maybe even more powerfully, imagination can directly impact both the episodic memories and semantic self-knowledge that comprise the self. Imagination changes the self by introducing false memories or changing existing memories. For example, after imagining themselves drinking soda, people both report that they drink more soda than they previously estimated, and they inflate their original estimate of how much soda they drink (Thomas, Hannula, & Loftus, 2007). Imagination can also change individuals' self-concept: imagining one's 'best possible self' increases trait ratings of optimism (Meevissen, Peters, & Alberts, 2011); imagining oneself voting changes one's attitudes towards voting (Libby, Valenti, Pfent, & Eibach, 2011); and imagining oneself in a stereotyped role increases self-stereotyping (Di Bella & Crisp, 2015). Purely internal processes can powerfully shape the self across a variety of domains.

One such internal process, simulation of other people, can also be a powerful source of change (Gilead et al., 2016). The term simulation can refer to multiple related concepts, from imagining counterfactual futures to thinking about other individuals or creating vivid mental imagery (Gaesser & Schacter, 2014; Markman, Gavanski, Sherman, & McMullen, 1993; Markman, Klein, & Suhr, 2012). In line with previous research on simulating other individuals (Mitchell, 2009; Meyer, Zhao, & Tamir, 2019; Nickerson, 1999), we operationalize simulation here as using aspects of one's own self-knowledge – whether episodic or trait knowledge – when thinking about other individuals. Spontaneous simulation occurs relatively frequently and is used for a variety of types of thought, including thinking about one's future or past, thinking about other places, hypothetical situations, or other individuals (Buckner & Carroll, 2007). Temporal simulation is also thought to underlie much of spontaneous thought (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009). In this manner, one's own self-knowledge and autobiographical memory serve as a source of social comparisons and inferences (Conway, Justice, & D'Argembeau, 2019; Gallese & Goldman, 1998; Schacter & Addis, 2007). This process, known as social simulation, can change self-knowledge (Meyer et al., 2019). More specifically, when individuals simulate another person (i.e., "My friend James likes dogs a lot"), they activate knowledge of their own self (i.e., "I like dogs a little bit"). This self-knowledge, once activated, becomes malleable. It can become so malleable, in fact, that it is susceptible to incorporating knowledge about the person being simulated. This results in one's self-concept shifting to become more similar to one's representation of the simulated individual (i.e., "I like dogs a lot"). This change in self-knowledge via simulation has been termed Simulation-Induced Malleability (SIM).

SIM has the potential to offer profound insight into how the self develops and transforms over time, though the full potential of this effect is as of yet unknown. Thus far, simulation has been shown to change both episodic memories and semantic self-knowledge (Meyer et al., 2019). We know that simulation can change episodic memories both

immediately and following a 24-h delay. To what extent are these changes durable even farther into the future? We also know that simulation can change trait knowledge, one feature of semantic self-knowledge. However, we do not know the extent to which these changes spread to semantically related features of the self. Here we test how deeply SIM can change self-knowledge in these two domains.

1.2. The durability of changes to self-knowledge

To what extent does SIM truly alter the content of self-knowledge? Here we test whether SIM causes merely momentary changes or lasting changes to autobiographical episodic memories. We propose that SIM might reflect at least two possible outcomes for the self: (i) Differential accessibility of self-knowledge during self-report, which would not induce lasting changes in self-knowledge, (ii) Sustained changes in the content or structure of self-knowledge, which would induce lasting changes in self-knowledge. Each of these possible outcomes derives from a distinct mechanism.

First, it is possible we see SIM because simulation changes what is active in working self-concept. According to research on working self-concept, different aspects of the self can be differentially active in different situations (Hinkley & Andersen, 1996; Markus & Wurf, 1987). That is, there is no single self-concept; rather, there is a "working self-concept" that is active in each moment and that can change depending on environmental and personal context. This working self-concept contains both core aspects of the self that are relatively stable, as well as aspects that are more malleable (Markus & Wurf, 1987). For example, a person might have a conception of herself as an employee, wife, and volunteer. When in the workplace, attributes of the employee-self may be heightened; a person may consider herself to be more professional and less emotive than she would when around her spouse. The self is thus a 'moving baseline' that can fluctuate in the short-term. This self should be resilient to short-term interventions, which would be unable to create a lasting change in self-knowledge (Demo, 1992). If simulation changes only working self-concept, then we would expect SIM to only bring about short-term change. That is, any effect of SIM should dissipate with sufficient time.

A more lasting outcome is predicted by work on memory malleability, more generally, and misinformation, more specifically. On these accounts, simulation alters the content of self-knowledge; in the same way, any kind of knowledge might be susceptible to change through misinformation. Misinformation effects occur when misleading or false information is presented after an event, and that new information updates existing episodic memories (Loftus, Miller, & Burns, 1978). For example, in a classic misinformation experiment, individuals are shown a video of a car driving through a red light. They are then exposed to false information, such as "the car was going fast when it drove through the *yellow* light." During a subsequent memory test, individuals are likely to believe that they saw the yellow light in the original video (Ayers & Reder, 1998; Loftus, 2005).

In classic misinformation experiments, changes to memory were still present when participants were questioned a week after the presentation of conflicting post-retrieval information (Loftus et al., 1978). Indeed, this effect has been shown to persist for as long as months to several years (Lommen, Engelhard, & van den Hout, 2013; Zhu et al., 2012). People may develop even more detailed, richer false memories over time, suggesting that the effect can be enhanced with time (Ost, Foster, Costall, & Bull, 2005; Schmolck, Buffalo, & Squire, 2000). While the question of how misinformation effects are instantiated has yet to be answered, it is clear that misinformation effects are long-lasting, if not permanent. If SIM acts as post-event information that alters an original memory, we would expect changes in memory to be long-lasting as well.

To arbitrate between these possible accounts of SIM, here we test the consequences of SIM on the durability of changes to self-knowledge. If SIM shifts working self-concept, simulation would not be sufficient to induce changes to more enduring aspects of self-knowledge. In this

scenario, we would expect to only see a short-term change to memory knowledge. On the other hand, if SIM truly modifies memories that are activated and made labile during simulation, SIM should persist for an extended period of time. The duration of the SIM effect can provide us with valuable insights into how deeply self-knowledge is altered by simulation.

1.3. The depth of changes to self-knowledge

Another question regarding SIM is the depth to which simulation changes semantic self-knowledge. Can simulating individuals on one aspect of self-knowledge result in changes to related traits? We propose that the answer to this will depend on whether simulation results in changes to self-knowledge that are deep enough to affect semantically related concepts. There are two possible outcomes: (i) Related traits will be changed as a result of simulation, or (ii) Only simulated traits will be changed by simulation; related traits will remain unaltered. Depending on which of these two outcomes we observe, we will be able to further elucidate how deeply simulation affects self-knowledge.

Self-knowledge consists of multiple components. In addition to autobiographical episodic memories, individuals also have beliefs and knowledge about themselves: i.e., knowing that they are allergic to kiwis, or thinking that they are generally a smart person. This self-knowledge is contained within conceptual self-knowledge and is stored in semantic memory, along with other general information (i.e., that the Earth is round, what the color 'blue' looks like; Conway et al., 2019; Conway & Bekerian, 1987; Klein, Loftus, & Kihlstrom, 1996). Early studies on SIM established that self-knowledge in the domain of personal semantic knowledge (i.e., traits) could be changed through simulation. To understand how deeply simulation may change semantic self-knowledge, it is useful to first outline relevant views on the structure of semantic knowledge.

One prominent class of semantic memory models are network models (Farah & McClelland, 1991; McRae, 2004). In network models, concepts are represented by individual nodes, each of which is connected to other nodes in a web-like organization (Collins & Loftus, 1975). Concepts that are semantically similar (i.e., 'sugar' and 'cake') are more tightly connected than concepts that are not semantically similar (i.e., 'sugar' and 'sofa'). Concepts that are more tightly connected are more likely to co-activate. For example, if a person is thinking about 'cake', a similar word, such as 'sugar', is more likely to be activated than 'sofa'. That is, 'sugar' is primed by 'cake', such that the person might be quicker to recognize the word than they would be to recognize 'sofa.' Experimental paradigms have demonstrated this effect through priming studies and memory retrieval studies (Hills, Jones, & Todd, 2012; Perea & Gotor, 1997; Perea & Rosa, 2002; Troyer, Moscovitch, & Winocur, 1997). Free association and mind wandering function similarly, such that thoughts flow along semantically constrained pathways (Gray et al., 2019; Mildner & Tamir, 2019; Polyn, Norman, & Kahana, 2009). Taken together, these findings indicate that closer semantic similarity begets closer association in memory recall (Collins & Loftus, 1975; Gruenewald & Lockhead, 1980; Hills et al., 2012; Saint-Aubin, Ouellette, & Poirier, 2005).

Semantic knowledge about the self is believed to be organized in the same manner as semantic knowledge more generally (Greenwald & Banaji, 1989). That is, semantically similar self-relevant concepts are linked in knowledge stores (Markus & Wurf, 1987). Research into self-concept structures suggests that the self-concept includes a complex network of interconnected traits, much like the network of semantic knowledge. Here, the traits included in the self-concept are not independent features, but rather conceptual nodes connected in a manner similar to the connectionist models of semantic knowledge (Elder, Cheung, Davis, & Hughes, 2020).

Semantic knowledge is conceptual in nature, and though it is accessed through language, it reflects knowledge that runs deeper than the words associated with it. Models of language learning and bilingualism hold that, while languages may use distinct words to represent a concept, the underlying semantic concept that is being accessed from each language is the same (Kroll & Stewart, 1994). These models distinguish between a lexical and conceptual level, with the notion of a deeper, conceptual level supported by cross-language priming studies. In these studies, for example, activation of one word (i.e. 'cat') will prime the same concept in another language (i.e., 'gato' in Spanish; Frenck & Pynte, 1987); translations of a word across languages thus access an underlying shared concept. That said, the words in each language can take on slightly different meanings due to the differing context and associations in each language. As such, translations represent semantically similar, though not identical, pairs of words.

This structure of semantic self-knowledge means that activating one piece of self-knowledge (i.e., 'I am 'kind'') might lead to activation of another close piece of self-knowledge (i.e., 'I am a 'compassionate''; Elder et al., 2020, Segal & Vella, 1990; Trafimow, Silverman, Fan, & Fun Law, 1997). If spreading models of activation capture the structure of semantic self-knowledge, then activating one trait should cause co-activation of semantically similar traits, and further, changing one trait through simulation should lead to changes in semantically similar traits. That is, when individuals simulate another individual on a specific trait, they will activate and change self-knowledge for that particular trait (i.e., 'kind'). This, in turn, should co-activate and change related traits (i.e. 'compassionate'). If simulation truly activates and changes semantic self-knowledge at a deep conceptual level, we should see the effects of SIM extend beyond the simulated trait, such that it affects associated traits as well. However, it is possible that this model does not accurately capture the structure of self-knowledge, that simulation only activates lexical knowledge, or momentarily-activated working knowledge, and not knowledge the conceptual level of a semantic structure, or simulation does not truly modify conceptual knowledge. If any of these possibilities hold, we would not expect to see any changes in self-knowledge for traits that are semantically related to the simulated trait. The extent to which simulation spreads will inform us about the depth of changes induced by simulation.

Here we test these possibilities in two ways. First, we test whether SIM changes only a directly simulated trait, or whether it also changes related traits within the same language. Second, we measure whether SIM changes only a directly simulated trait, or whether it also changes self-knowledge of the same trait, translated into a different language. If SIM causes deep changes to semantic knowledge about the self, we expect it to alter both semantically related traits within a language, as well as translated traits across languages.

1.4. Current studies

The SIM effect shows that the self is not static, and instead is subject to the forces of change caused by internal simulation. Initial research on SIM established that it could change one's positive and negative traits and episodic memories, that it could be measured via self-report and naturalistic language, and that it could last for 24 h. However, the extent to which this malleability changes self-concept has yet to be explored. The present research aims to establish the durability and depth to which SIM alters self-knowledge in several ways.

First, we explore the durability of SIM. In Study 1, we aim to replicate the SIM effect, testing whether simulating others can change the valence of episodic memories to be more similar to the other at zero delay. In Study 2a, we aim to replicate previous findings demonstrating that SIM can persist after a 24-h delay. In Study 2b, we test for SIM persisting after imposing a delay of 48-h. Second, we explore the depth of SIM by measuring how far its influence spreads through semantic self-knowledge. In Study 3, we test whether simulation of a trait will cause changes in self-ratings of a semantically related trait; in Study 4, we formally calculate changes across both highly and moderately semantically similar traits. In Study 5, we test whether this effect will hold when the trait pairs are cross-language translations. Together, these

studies explore the extent to which SIM can change our episodic memories and semantic self-knowledge.

2. Study 1: Establishing SIM

2.1. Method

The goal of Study 1 is to establish that simulation can indeed make self-knowledge malleable. That is, we will replicate prior work demonstrating SIM, operationalized here as changes in participant self-ratings such that they become more similar to their ratings of a simulated target. Participants first considered a series of their own episodic memories, and rated how they felt during those events. They then simulated another person experiencing similar events, and rated their experiences. Finally, they re-rated how they would feel in the same set of memories as in the first stage. This allowed us to explore how one's own memories change after simulating others in a similar situation. The simulated targets were either a similar other or a dissimilar other, allowing us to test if this change in memories depends upon participant similarity to the target.

We report how we determined our sample sizes, all a priori data inclusion/exclusion criteria, all manipulations, and all measures in this study and all subsequent studies. In all studies, we maximized power by determining sample sizes after conducting power analyses on pilot studies or prior studies, using the crowdsourcing platform Prolific instead of Amazon Mechanical Turk to ensure higher-quality data, and instituting complete randomization of stimuli within and across conditions. Data and analysis code for this and all subsequent studies can be accessed here: https://osf.io/ga65b/?view_only=8d780eb06d0e4f42a62f2712825c4f6e.

2.2. Participants

Participants ($N = 186$) were recruited from Prolific. This sample size provided at least 99% power to detect a target x time interaction effect, at a size reported by prior work on this same effect ($\eta_p^2 = 0.05$; Meyer et al., 2019). Participants were excluded prior to analyses based on three a priori exclusion criteria: if they completed the task in an unreasonably short period of time ($n = 2$), if they self-reported poor English proficiency, or if they provided fewer than 15 unique answers. These exclusions left a final sample size of 184 participants (mean age = 39.2, 134 female). Participants in this and all subsequent studies provided informed consent in accordance with the Princeton University Institutional Review Board.

2.3. Procedure

There were three main phases in this experiment (Fig. 1). In the first *baseline phase*, participants were asked to recall a series of memories. Each of the prompts was intended to probe a specific memory, such as "remember a time you received an award." There were 42 memory prompts, half positive and half negative. After recalling each memory, participants rated how they felt during that experience on a scale from 'Extremely Negative' to 'Extremely Positive.' This scale was coded from 1 to 100, though numbers were not displayed to participants to prevent any numerical anchoring of their responses.

In the second *simulation phase* of the experiment, participants were asked to imagine another individual in hypothetical versions of the same scenarios they had just imagined themselves in. Participants imagined two targets, with the order counterbalanced across participants. For one target, participants were asked to provide the name of a friend who they felt had similar personality, temperament, major likes and dislikes, beliefs, and values. This friend's name was then integrated into the questions; i.e., "How would Josh feel receiving an award?" For the other target, participants simulated the average American to reflect a dissimilar target.

Participants simulated each target on 14 simulation prompts (7 positive and 7 negative), each of which provided a hypothetical version of an episodic memory prompt seen in the baseline phase. Prompts were randomly paired with only one of the simulated targets for each participant. Participants used the same sliding scale of 'Extremely Negative' to 'Extremely Positive' to rate how each target would feel in each situation.

In the *post-simulation phase*, participants again rated how they felt
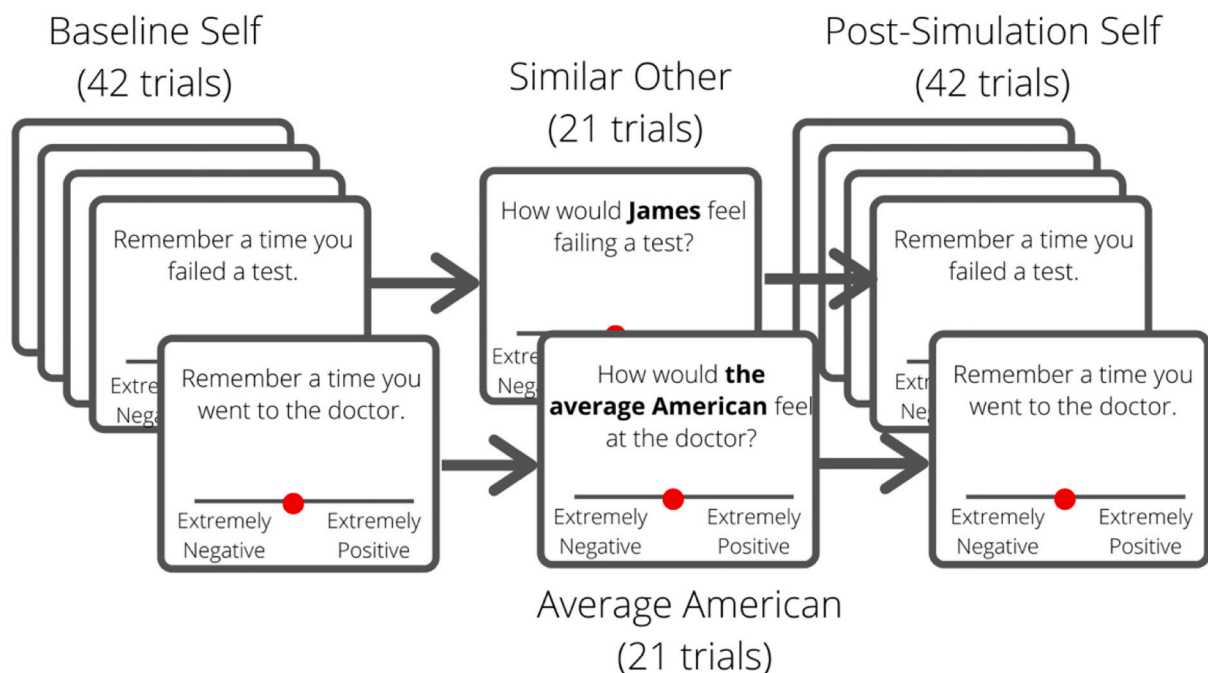
Fig. 1. Task Schematic for Study 1. Participants began by rating 42 of their episodic memories, followed by simulating a similar other and the average American on a subset of the initial 42 memories, with the order counterbalanced across participants. Participants then re-rated themselves on the 42 initial memories.

during each of the original episodic memories. All 42 prompts were presented in a randomized order.

Following the three main phases of the experiment, participants were asked to provide ratings on each of the targets. To confirm that participants felt more similar to their friend than to the average American, participants were asked to rate how similar they felt to each target. Participants also reported how emotionally close they felt to their friend and categorized their relationship with their friends, measures which were not analyzed as part of this and all subsequent future studies.

According to previous literature on SIM, we expected simulation to change individuals' self-knowledge for each prompt from baseline to post-simulation, such that it becomes more like that of the simulated target. This would appear in our analyses as a main effect of time. We expected to see a simple effect of time for both targets, such that simulation changes self-ratings to be more like the target regardless of which target is being simulated. If we see a main effect of time, we would have reason to believe that the self-knowledge that is accessed while simulating others is indeed prone to misinformation effects. In addition, we expected to see an effect of target, as participants should rate themselves more similarly to a similar target than the average American. Finally, we expected to see an interaction between time and target, such that individuals' self-ratings will change by a greater amount when the target simulated is similar than when the target is dissimilar.

2.4. Results

To confirm that our similarity manipulation was successful, we compared individuals' self-reported similarity to the similar other and the average American in a paired t-test. The t-test showed that participants felt more similar to the self-generated similar target (M = 80.1, SD = 16.1) than to the average American (M = 34.3, SD = 22.6). This confirms that our manipulation of similarity between the targets was successful (t(182) = 24.9, p < .001, Cohen's d = 2.34).

The SIM effect is defined here as individuals' memories or traits changing to become more like another individual. We operationalize evidence for SIM in this paradigm as self-ratings shifting after simulation to become like those of a simulated target. To explore this effect, we calculated two values for each prompt: the *baseline difference*, which refers to the absolute difference between the self-ratings at the baseline phase and the target ratings from the simulation phase, and the *post-simulation difference*, which refers to the absolute difference between self-ratings in the post-simulation phase and the target ratings from the simulation phase. A smaller difference score means that the participants' self-ratings are closer to the target ratings. The baseline and post-simulation difference scores are first calculated for each independent item, and the difference values are then averaged across all prompts of each target. This provided us with four values: the baseline difference for prompts associated with the similar other; the baseline difference for prompts associated with the average American; the post-simulation difference for prompts associated with the similar other; and the post-simulation difference for prompts associated with the average American.

We tested for SIM using a 2 (time: baseline and post-simulation) x 2 (target: average American and similar other) repeated-measures ANOVA (Fig. 2). As expected, this ANOVA revealed a significant main effect of time (F(1,183) = 69.3, η_p^2 = 0.275, p < .001), such that participants' self-ratings were more similar to the target ratings at the post-simulation period (M = 10.93, SD = 5.15) than at baseline (baseline M = 12.65, SD = 5.75). This replicates prior work on SIM by demonstrating that simulating a target can cause episodic memories to shift towards knowledge of the simulated target. This ANOVA also revealed the expected main effect of target (F(1,183) = 10.6, η_p^2 = 0.055, p < .001), such that participant self-ratings were more similar to the similar other (M = 11.16, SD = 5.55) than the average American (M = 12.42, SD = 5.43). Contrary to expectations, no interaction between time and target was observed (F(1,183) = 0.033, η_p^2 = 1.8e-4, p = .86).

Finally, we tested for a simple effect of time within each target
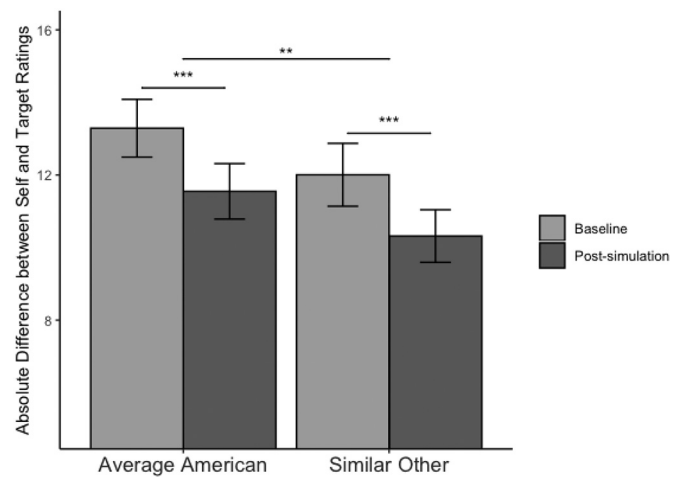
Fig. 2. Y-axis represents the difference between self and target ratings; smaller values reflect greater similarity between self and target. Light gray bars reflect similarity before simulation, while dark gray bars reflect similarity after simulation. X-axis represents the target simulated. Error bars represent 95% confidence intervals; * = p < .05; ** = p < .01; *** = p < .001.

(average American and similar other). We observed significant changes in self-knowledge from baseline to post-simulation for both the average American (post-simulation M = 11.55, SD = 4.26; baseline M = 13.29, SD = 5.47) and the similar other (post-simulation M = 10.32, SD = 4.98; baseline M = 12.01, SD = 5.95); the effect size was similar for both targets (average American: (t(183) = 7.66, p < .001, Cohen's d = 0.32; similar other: (t(183) = 6.27, p < .001, Cohen's d = 0.31).

These results show that after simulating another individual in specific circumstances, people change their self-ratings on episodic memories to be more like that of the target they simulated. This establishes the SIM effect demonstrated in prior work, supporting the notion that the self can be changed through simulation.

3. Study 2: Temporal duration

3.1. Study 2a: 24-h delay

3.1.1. Method

This study was intended to replicate earlier findings that SIM can be observed 24 h after simulation (Meyer et al., 2019). If such an effect is reliably observed, we would have further support for the hypothesis that simulating another individual changes episodic self-knowledge on a more than superficial level. A main effect of time after a delay period would provide evidence that simulating another individual can serve as post-event information, altering individuals' episodic memories.

3.1.2. Participants

Participants (N = 185) were recruited from Prolific. This sample size provides 99% power to detect the main effect of time observed in Study 1. Participants were excluded prior to analyses based on two a priori exclusion criteria: if they completed the study in an unreasonably short period of time or if they provided fewer than 15 unique answers (n = 1). In addition, 29 participants did not return for the second phase, and as such, were removed from analyses. These exclusions left a final sample size of 152 participants (mean age = 37.18, 80 female).

3.1.3. Procedure

The study was identical to Study 1, with the addition of a 24-h delay between the simulation and the post-simulation phases. Participants rated themselves on 42 memories and then simulated an average American on a subset of 14 questions and a self-generated similar other on a different subset of 14 questions. The post-simulation phase

occurred after a 24-h delay, when they then re-rated themselves on the same 42 initial situations.

3.2. Results

We again performed a paired *t*-test of participants' similarity ratings to the similar other and average American. Of the 152 participants included in the analysis, 102 provided ratings of psychological similarity to both targets. As in Study 1, a t-test confirmed our similarity manipulation, such that individuals reported greater psychological similarity to the similar other (M = 77.76, SD = 14.33) than to the dissimilar other (M = 39.59, SD = 19.03; t(101) = 18.12, $p < .001$, Cohen's d = 2.27).

As in Study 1, we tested for SIM with a 2 (time: baseline and post-simulation) x 2 (target: average American and similar other) repeated-measures ANOVA (Fig. 3). Results fully replicated the results of Study 1. This ANOVA revealed a significant main effect of time (F(1,151) = 10.1, $\eta_p^2 = 0.063$, $p = .0018$), such that participant self-ratings were more similar to the target ratings post-simulation (M = 12.59, SD = 6.07) than at baseline (M = 13.24, SD = 5.87). This result replicates the findings from a previous study that SIM persists even after a 24-h delay between the simulation and post-simulation phases. This ANOVA also revealed the expected main effect of target (F(1,151) = 19.8, $\eta_p^2 = 0.16$, $p < .001$), such that participants' self-ratings were more similar to the similar other (M = 12.03, SD = 5.62) than the average American (M = 13.80, SD = 6.20). As in Study 1, but unlike previous research with a 24-h delay, we observed no interaction between time and target (F(1,151) = 0.14, $\eta_p^2 = 9.4e-4$, $p = .71$).

The simple effect of time within each target (average American and similar other) again revealed a significant effect of time for both targets. The effect size was similar for both the average American (t(151) = 2.32, $p = .022$, Cohen's d = 0.094; post-simulation M = 13.51, SD = 6.45; baseline M = 14.09, SD = 5.95) and the similar other (t(151) = 2.62, $p < .01$, Cohen's d = 0.13; post-simulation M = 11.68, SD = 5.55; baseline M = 12.39, SD = 5.68).

These results support previous evidence and suggest that SIM persists for at least 24-h for episodic memories, albeit to a lesser extent than after no delay. That is, after simulating a target, self-knowledge a day later remains shifted towards that target. This suggests that, when individuals simulate another person, the simulated event can retroactively interfere

with existing episodic memories. These findings help us understand whether SIM involves altering episodic memories in a lasting manner.

3.3. Study 2b. 48-h delay

3.3.1. Method

Studies 1 and 2a replicate earlier findings that the SIM effect – characterized as increased similarity between individual and target ratings after simulation – appears immediately and persists even after a 24-h delay for episodic memories. Study 2b tests whether SIM persists at least 48 h after simulation. If SIM persists 48 h after memory reactivation, strong credence would be lent to the notion that personal memories are altered by simulation.

3.3.2. Participants

Participants ($N = 450$) were recruited from Prolific. This sample size provides 99% power to detect the main effect of time observed in Study 2a. Participants were excluded prior to analyses based on three a priori exclusion criteria: if they completed the task in an unreasonably short period of time ($n = 1$), if they self-reported poor English proficiency ($n = 3$), or if they provided fewer than 15 unique answers. An additional 6 participants were removed for not providing identification numbers, leaving us unable to match their data from the first and second phases of the study, and 2 participants were removed for failing to answer all of the questions. In addition, 65 participants did not return for the post-simulation phase, and as such, were removed from analyses. These exclusions left a final sample size of 373 participants (mean age = 38.37, female = 263).

3.3.3. Procedure

Study 2b used the same procedure as in Study 2a, with a 48-h delay instead of a 24-h delay. Participants completed the baseline phase by providing initial self-ratings on 42 memories and simulated an average American on 1/3 of these stimuli and a self-generated similar other on another 1/3 of the stimuli. Following a 48-h delay, participants completed the post-simulation phase by rating the valence of the original episodic memories again. In order to best replicate previous studies on SIM, perceived similarity to the target was recorded after the second day of participation.
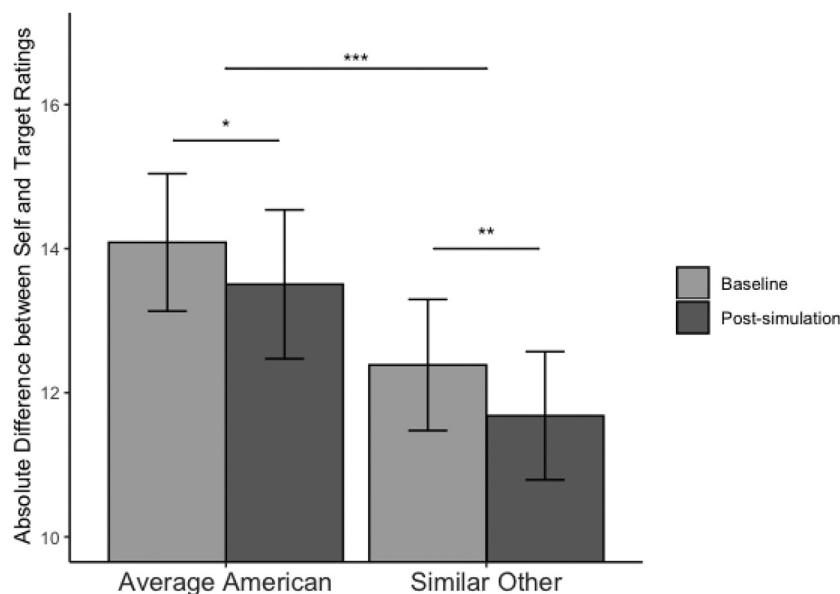

Fig. 3. Y-axis represents the difference between self and target ratings; smaller values reflect greater similarity between self and target. Light gray bars reflect similarity before simulation, while dark gray bars reflect similarity after simulation. X-axis represents the target simulated. Error bars represent 95% confidence intervals; * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

3.4. Results

We performed a paired *t*-test of participants' similarity ratings to the similar other and average American. This t-test of target on reported similarity again confirmed our similarity manipulation, such that individuals reported greater similarity to the similar other (M = 78.54, SD = 19.25) than to the dissimilar other (M = 43.51, SD = 20.53; t(372) = 26.18, *p* < .001, Cohen's d = 1.76).

As in Study 1, we tested for SIM with a 2 (time: baseline and post-simulation) by 2 (target: similar other or average American) repeated-measures ANOVA (Fig. 4). This ANOVA revealed a main effect of time (F(1,372) = 4.33, η_p^2 = 0.011, *p* = .038), signifying that participants' post-simulation ratings (M = 12.49, SD = 5.58) were significantly closer to the target ratings than their baseline simulation ratings (M = 12.79, SD = 5.52). The expected significant effect of target was also observed (F (1,372) = 8.59, η_p^2 = 0.023, p < .001), such that participants rated themselves as more similar to the similar other (M = 12.26, SD = 5.32) than to the average American (M = 13.02, SD = 5.74),). No significant interaction between these variables was observed (F(1,372) = 0.52, η_p^2 = 0.0014, *p* = .47).

We tested for a simple effect of time within each target. Despite observing a main effect of time, we observed no significant simple effects reflecting changes in self-knowledge towards the target for either the average American (t(372) = 1.96, *p* = .05, Cohen's d = 0.14; post-simulation M = 12.83, SD = 5.85; baseline M = 13.22, SD = 5.63) or the similar other (t(372) = 1.21, *p* = .23, Cohen's d = 0.089; post-simulation M = 12.15, SD = 5.28; baseline M = 12.37, SD = 5.37).

The results from Studies 2a and 2b show that SIM can be observed even with an extended delay between simulation and testing. These data suggest that simulating another individual can cause changes in personal episodic memories for an extended period of time, a finding with many implications for future research on and potential uses for SIM.

4. Study 3: Depth of change: synonyms

4.1. Method

Studies 1–2 show that when simulation involves episodic memories, simulation of another individual can be incorporated into episodic memories, altering self-knowledge beyond what working self-concept can account for. Though episodic and semantic self-knowledge are distinct, prior work on SIM suggests that trait-related self-knowledge is vulnerable to change based on simulation, just like other types of knowledge (Klein, Chan, & Loftus, 1999; Klein & Lax, 2010; Meyer et al., 2019). When participants simulate another individual by rating how well traits apply to them, their ratings of how well those traits apply to themselves change in the direction of the target. This effect occurred for both positive and negative traits and when the target was a similar or a dissimilar other.

This previous work showed that we can change self-ratings on a trait after simulating another individual on that trait. We are curious as to the extent of this change in self-knowledge. Is it possible for simulating an individual on one trait to change self-knowledge on a semantically close related trait? To test if SIM can change self-knowledge on semantically similar concepts, Study 3 adapts the initial trait paradigm to test the generalizability of SIM to synonymous adjectives. If SIM generalizes to the extent of affecting semantically similar concepts, we would expect to see participants' self-ratings on one trait become more like their ratings of a synonymous but not identical trait. If not, we would see the presence of SIM for identical traits only. We preregistered our hypotheses and analyses for Study 3 through the Open Science Framework: https://osf. io/t8bg2?view_only=6a9447ecd6f4494da2f8859aa9475a72.

4.1.1. Participants

Participants (N = 196) were recruited from Mechanical Turk. The sample size was estimated to detect the main effect of time observed in Study 1 at 99% power. Participants were excluded prior to analyses based on two a priori exclusion criteria: if they completed the task in an unreasonably short period of time (*n* = 5) or if they provided fewer than 10 unique answers (*n* = 6). These exclusions left a final sample size of 185 participants (59 female, 102 male, 24 unavailable).

4.1.2. Procedure

There were three phases in this experiment (Fig. 5). In the first *baseline phase*, participants began by rating themselves on 28 traits. Traits were selected to be matched in familiarity, frequency of use, and valence (Dumas, Johnson, & Lynch, 2002). Half were positive in valence, and half were negative in valence. Stimuli were selected from a
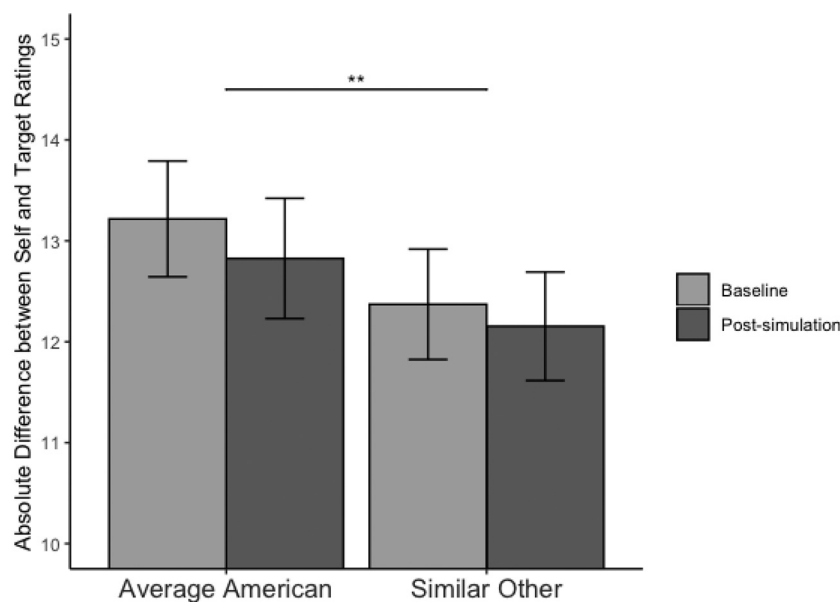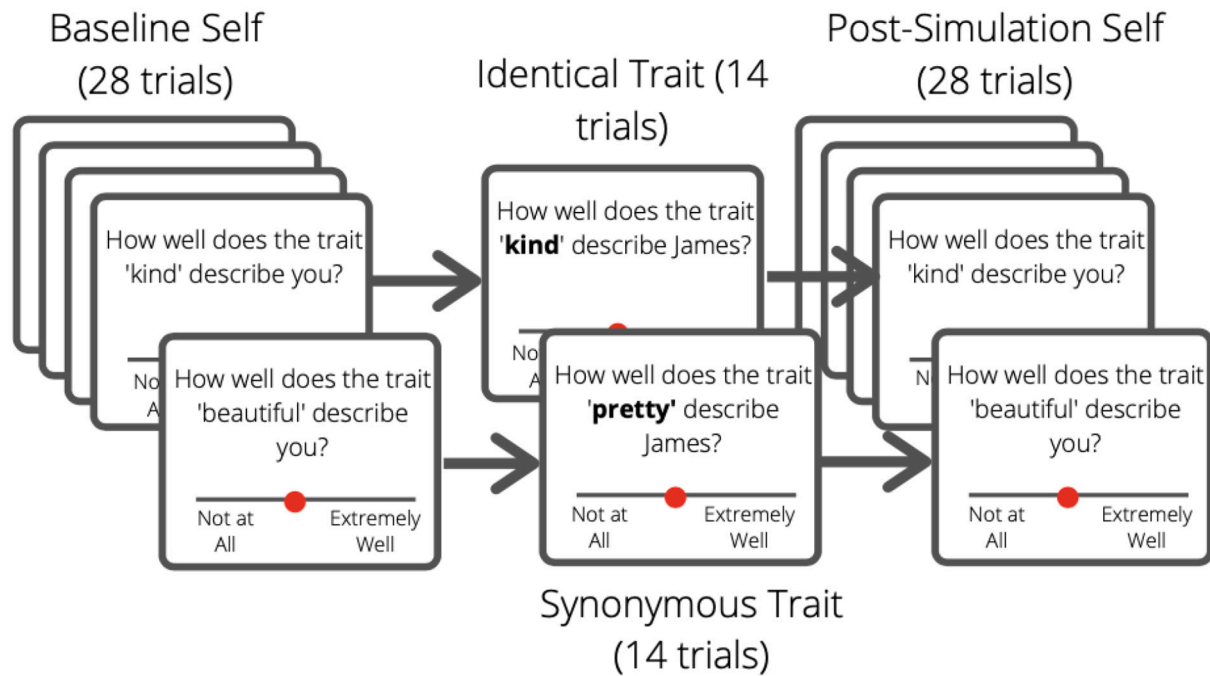

Fig. 4. Y-axis represents the difference between self and target ratings; smaller values reflect greater similarity between self and target. Light gray bars reflect similarity before simulation, while dark gray bars reflect similarity after simulation. X-axis represents the target simulated. Error bars represent 95% confidence intervals; * = *p* < .05; ** = *p* < .01; *** = *p* < .001.

Fig. 5. Task Schematic for Study 3. Participants began by rating themselves on 28 traits. They then rated a similar other on synonyms of half of the original set of traits and the identical trait of the other half. These ratings for identical and synonymous traits were randomly presented within one block. Participants then re-rated themselves on the 28 initial traits.

list of 28 pairs of synonymous words (e.g., intelligent & smart, kind & compassionate), with one word from each pair randomly selected to be presented in the baseline phase. Trait pairs were generated using WordNet synsets (Miller, 1995). For a complete list of trait pairs, please see the Supplement. The traits were presented in a random order to participants, who rated how well the trait applied to them on a continuous scale, from 'Not at all' to 'Extremely well.' The continuous scale corresponded to a numerical 1–100 scale, though numerical values were not visible to participants.

In the second *simulation phase*, participants then provided the name of a similar friend and were asked to rate how well 28 traits applied to this friend. Participants rated their friend on two types of traits: identical or synonymous. Half of the traits were identical to the ones were presented during the baseline self-ratings phase (e.g., 'intelligent' for self, and 'intelligent' for the friend). The other half were synonymous with the traits presented during the baseline phase (e.g., 'kind' for the self and 'compassionate' for the friend).

In the *post-simulation phase*, participants were then asked to rate themselves again on the same 28 words they rated themselves on initially. Finally, participants were asked to rank their similarity to their
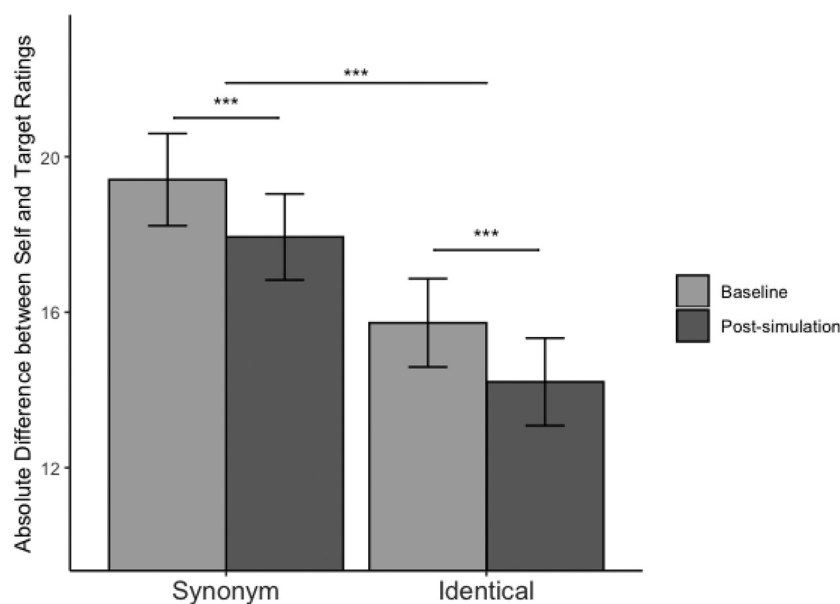
Fig. 6. Y-axis represents the difference between self and target ratings; smaller values reflect greater similarity between self and target. Light gray bars reflect similarity before simulation, while dark gray bars reflect similarity after simulation. X-axis represents the type of word simulated. Error bars represent 95% confidence intervals; * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

friend.

If simulation co-activates and makes semantically related traits labile, we would expect to see a main effect of time. Importantly, we would also expect a *t*-test to reveal a simple effect of time within both the identical and the synonymous conditions.

4.2. Results

We tested for SIM using a 2 (time: baseline and post-simulation) x 2 (type of word simulated: synonyms or identical words) repeated-measures ANOVA (Fig. 6). Results support the hypothesis that simulating a target changes trait-based knowledge. The ANOVA revealed a significant main effect of time ($F(1,184) = 56.43$, $\eta_p^2 = 0.23$, $p < .001$), such that participants' self-ratings were closer to the target ratings in the post-simulation phase ($M = 16.07$, $SD = 7.91$) than in the baseline phase ($M = 17.57$, $SD = 8.20$). A main effect of the type of word simulated was also observed ($F(1,184) = 50.1$, $\eta_p^2 = 0.21$, $p < .001$), with the identical words ($M = 14.97$, $SD = 7.82$) showing a smaller difference between the self and the target than the synonyms ($M = 18.68$, $SD = 7.93$). No interaction between these variables was observed ($F(1,184) = 0.016$, $\eta_p^2 = 8.5\text{e-}5$, $p = .90$), suggesting that the effect for identical traits was not different from that of synonymous traits.

Additionally, the simple effect of time was calculated within each type of word (synonymous and identical words). We observe significant changes in self-knowledge towards the target for both types of word; the effect size was similar for the synonymous words ($t(184) = 5.46$, $p < .001$, Cohen's $d = 0.19$; post-simulation $M = 17.94$, $SD = 7.64$; baseline $M = 19.41$, $SD = 8.17$) and identical words ($t(184) = 5.68$, $p < .001$, Cohen's $d = 0.19$; post-simulation $M = 14.21$, $SD = 7.76$; baseline $M = 15.72$, $SD = 7.84$).

Together, these results show that SIM can be observed when not only identical words or situations are simulated, but also when using synonyms. This suggests the activation of self-knowledge when simulating others goes beyond pure visual representation and helps further the argument that, at least with regards to semantic knowledge, SIM is not merely based on working self-concept but affects deeper self-knowledge.

5. Study 4: depth of change: semantic spread

5.1. Method

Study 3 showed that individuals change self-ratings on a trait after simulating another individual on synonymous traits. While the previous study used synonymous traits selected from a dictionary and WordNet, here we looked to formalize the semantic relationship between simulated traits. We further sought to explore whether the degree of semantic similarity impacts the extent to which self-knowledge is changed. Study 4 adapts the paradigm used in Study 3, but with the semantic relationship between traits formalized as a distance in semantic space. Simulated traits were selected to be either highly similar, moderately similar, or not at all similar to a set of baseline traits. If simulation affects both highly and moderately similar traits, then we would expect to see that participant's self-ratings become more like the target, as in Study 3. We preregistered hypotheses and analyses for Study 4 through the Open Science Framework: https://osf.io/ah8gz?view_only=a61877854b2949309aaad76c94026504.

5.1.1. Participants

Participants ($N = 200$) were recruited from Prolific. The sample size was estimated to detect the main effect of time observed in Study 3 at 99% power. Participants were excluded prior to analyses based on three a priori exclusion criteria: if they completed the task in an unreasonably short period of time ($n = 1$), if they reported poor English proficiency, or if they provided fewer than 7 unique answers. These exclusions left a final sample size of 199 participants (135 female, 62 male, 1 nonbinary, 1 other).

5.1.2. Procedure

There were three phases in this experiment (Fig. 7). In the first *baseline phase*, participants rated themselves on 15 base traits. The traits were presented in a random order, and participants rated how well each trait applied to them on a continuous scale, from 'Not at all' to 'Extremely well.' The continuous scale corresponded to an unmarked numerical 1–100 scale.

In the second *simulation phase*, participants then provided the name of a similar other and were asked to rate how well 15 traits applied to this friend. Participants rated the similar other on three types of traits: highly similar traits, moderately similar traits, or control traits. A third of the traits were highly similar to the ones were presented during the baseline self-ratings phase (e.g., 'funny' for self and 'amusing' for the similar other). Another third were moderately similar to the base traits (e.g., 'funny' for the self and 'witty' for the similar other). The final third were control traits, which referred to physical characteristics (e.g. 'funny' for the self and 'clean-shaven' for the similar other).

In the *post-simulation phase,* participants were then asked to rate themselves again on the 15 base traits they rated themselves on initially. Finally, participants were asked to rank their similarity to the similar other and to provide demographic information.

The initial 15 base traits were selected from a list of personality traits, pre-rated for likeableness and meaningfulness; of these, 7 were extremely likable, 7 were extremely unlikeable, and 1 was neutral (Gray et al., 2019). For each of these traits, we generated semantically similar matches based on the semantic similarity between the base trait and the match. We performed this calculation by taking the location of each trait in a multidimensional semantic space using fastText, with vectors pre-trained on Common Crawl and Wikipedia (Bojanowski, Grave, Joulin, & Mikolov, 2017; Joulin, Bojanowski, Mikolov, Jegou, & Grave, 2018). We extracted the vector corresponding to the location of each trait in semantic space and calculated semantic similarity between each pair of traits as the cosine between the vectors representing each word. This allowed us to select 15 triads of traits (45 traits total), with one base trait, one highly similar trait, and one moderately similar trait. All 45 traits were compared in a matrix to confirm that no trait in a triad would interfere with traits in another triad; that is, even the moderately similar trait was more closely correlated to its associated base trait than it was to any other trait. The control traits were selected from a list of normed traits describing physical characteristics and were randomly assigned to base traits. For a complete list of traits, please see the Supplement.

Based on the results observed in Study 3, we expected simulation to make semantically related traits labile; this should be reflected in a main effect of time. In all cases, we expected an interaction effect, such that individuals' self-ratings shift to become more like the target for both highly and moderately similar traits, but not for the control traits. Note, the changes in similarity between the highly and moderately similar trait were minimal. As such, we did not expect an interaction in the strength of the effect between these two conditions. Finally, we anticipated a simple effect of time in both experimental conditions and not in the control condition.

5.2. Results

We tested for SIM using a 2 (time: baseline and post-simulation) x 3 (type of word simulated: control, higher-similarity, or lower-similarity) repeated-measures ANOVA (Fig. 8). Results support the hypothesis that simulating a target on semantically related characteristics changes trait-based knowledge. The ANOVA revealed a significant main effect of time ($F(1,196) = 18.91$, $\eta_p^2 = 0.087$, $p < .001$), such that participants' self-ratings were closer to the target ratings in the post-simulation phase ($M = 26.35$, $SD = 13.78$) than in the baseline phase ($M = 27.39$, $SD = 13.07$). As expected, we saw an interaction between time and condition ($F(2,396) = 5.66$, $\eta_p^2 = 7.45\text{e-}4$, $p = .0038$), driven by the expected SIM effect for the highly and moderately similar words, and a lack of SIM effect in the control condition. In addition, we observed a main effect of
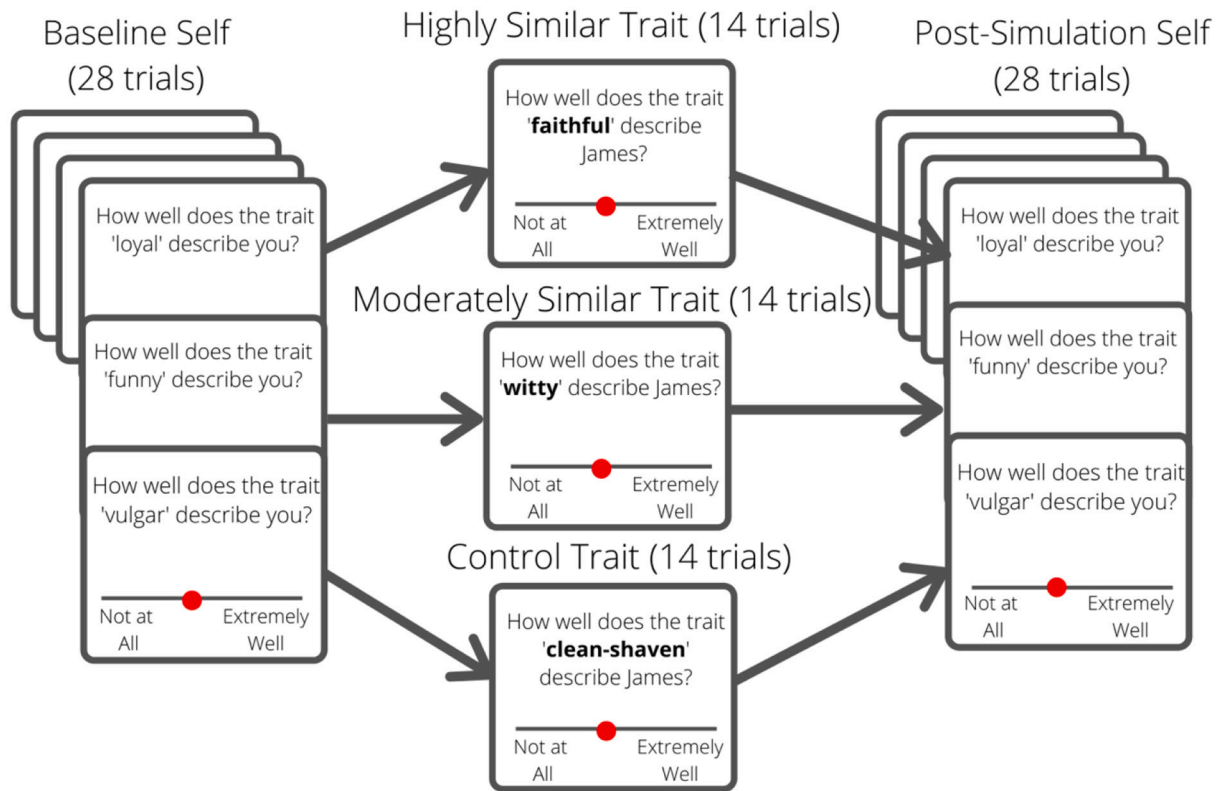
Fig. 7. Task Schematic for Study 4. Participants began by rating themselves on 15 base traits. They then rated a similar other on 5 words that were highly similar to the base traits, 5 words that were moderately similar to the base trait, and 5 words that were unrelated and described physical characteristics. These ratings were presented in a random order within one block. Participants then re-rated themselves on the 15 base traits.
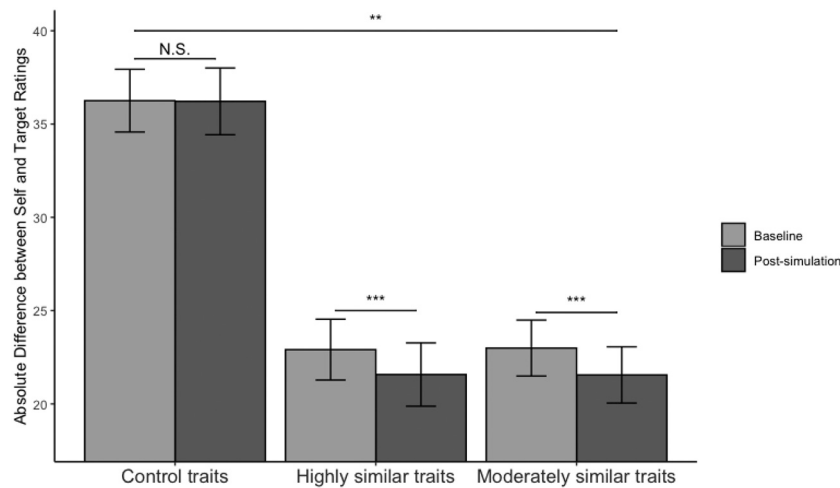
Fig. 8. Y-axis represents the difference between self and target ratings; smaller values reflect greater similarity between self and target. Light gray bars reflect similarity before simulation, while dark gray bars reflect similarity after simulation. X-axis represents the type of trait simulated. Error bars represent 95% confidence intervals; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; N.S. = Not Significant.

the type of word simulated was also observed ($F(2,396) = 126.7$, $\eta_p^2 = 0.24$, $p < .001$; higher-similarity words M = 22.24, SD = 11.90; lower-similarity words M = 22.27, SD = 10.77), with the effect driven by the control condition (M = 36.23, SD = 12.40).

Additionally, the simple effect of time was calculated within each condition. We observe significant changes in self-knowledge towards the target for both experimental conditions; the effect size was similar for the highly similar traits ($t(198) = 3.88$, $p < .001$, Cohen's d = 0.11; post-simulation M = 21.57, SD = 12.14; baseline M = 22.9, SD = 11.65) and

moderately similar traits ($t(198) = 4.46$, $p < .001$, Cohen's d = 0.13; post-simulation M = 21.55, SD = 10.78; baseline M = 22.99, SD = 10.73). As expected, no effect of time was observed in the control condition ($t(198) = 0.10$, $p = .92$, Cohen's d = 0.003; post-simulation M = 36.22, SD = 12.80; baseline M = 36.25, SD = 12.02).

Taken together with Study 3, these results show that SIM can extend to affect not only identical but semantically related concepts. This relationship persists for both highly semantically related as well as moderately semantically related traits, suggesting that the SIM effect is

capable of spreading through semantic similarity networks. More broadly, these results indicate that, with regards to trait knowledge, simulating other individuals is capable of changing self-concept to an extensive degree.

6. Study 5: depth of change: cross-language

6.1. Method

Studies 3–4 show that we can change self-ratings on a trait after simulating another individual on semantically related traits. Synonyms represent one such example of semantically related traits. Here, we test a second example of semantically related traits by using cross-language translations. We ask: does simulating an individual on one trait change self-knowledge on a translation of that trait? To test this question, Study 5 adapts the paradigm from Study 3. On half the trials, participants still rate themselves and a target on an identical trait word. On the other half of the trials, participants make these ratings on translated pairs of traits. If SIM affects semantic self-knowledge at the level of concepts rather than just words, we would expect to see participants' self-ratings on one characteristic become more like their ratings of the same trait in another language. If not, we would see the SIM effect present only for the identical traits. We preregistered our hypotheses and analyses for Study 5 through the Open Science Framework: https://osf.io/dpafe?view_only=8c4dc968409047e4977c30bfe4 d88270.

6.1.1. Participants

Participants ($N = 199$) who were bilingual in English and Spanish were recruited from Prolific. The sample size was estimated to detect the main effect of time observed in Study 3 at 99% power. Participants were excluded prior to analyses based on three a priori exclusion criteria: if they completed the task in an unreasonably short period of time ($n = 0$), if they provided fewer than 10 unique answers (n = 0), or if they reported themselves as not fluent in English and Spanish ($n = 27$). These exclusions left a final sample size of 172 participants (102 female, 69 male, 1 prefer not to respond).

6.1.2. Procedure

Participants were randomly assigned to one of two versions of the survey. In each version, one language (Spanish or English) was used as the *baseline language*, and all instructions were presented in that language. There were three phases in this experiment (Fig. 9). In the first *baseline phase*, participants began by rating themselves on 28 traits presented in the baseline language. The traits were presented in a random order to participants, who rated how well the trait applied to them on a continuous scale, from 'Not at all' to 'Extremely well,' which corresponded to a 1–100 unmarked scale.

In the second *simulation phase*, participants provided the name of a similar friend and were asked to rate how well 28 traits applied to this friend. Participants rated their friend on two types of traits: same-language traits or translated traits. Half of the traits were presented in the baseline language (the same-language traits; these traits were identical, e.g., 'beautiful' for the self and 'beautiful' for the friend), while the other half were presented in the other language of fluency (the translated traits; e.g. 'beautiful' for the self and 'hermoso' for the friend). These traits were blocked, and the order of the blocks was randomized between participants.

In the *post-simulation phase*, participants were then asked to rate themselves again on the 28 words they rated themselves on initially, presented in the baseline language. Finally, participants answered a series of questions about their relationship to the friend and their bilingual status.

The 28 traits were selected from an initial list of 75 traits, which were the most frequent English personality adjectives rank-listed in the COCA corpus (Davies, 2008). Each of the 75 traits was translated into Spanish, with cognate translation pairs (i.e. 'efficient'-'eficiente') removed. Pairs where the Spanish translations were present in the Corpus del Español (Davies, 2016) fewer than 10,000 times were removed. This was done to ensure that traits were of relatively equal frequency in their respective languages. The location of each trait in a high-dimensional semantic space was calculated using fastText, with vectors trained on Common Crawl and Wikipedia (Bojanowski et al., 2017; Joulin et al., 2018). The similarity between traits across languages was determined by taking the cosine of the two vectors. The 14 positive and 14 negative traits with the
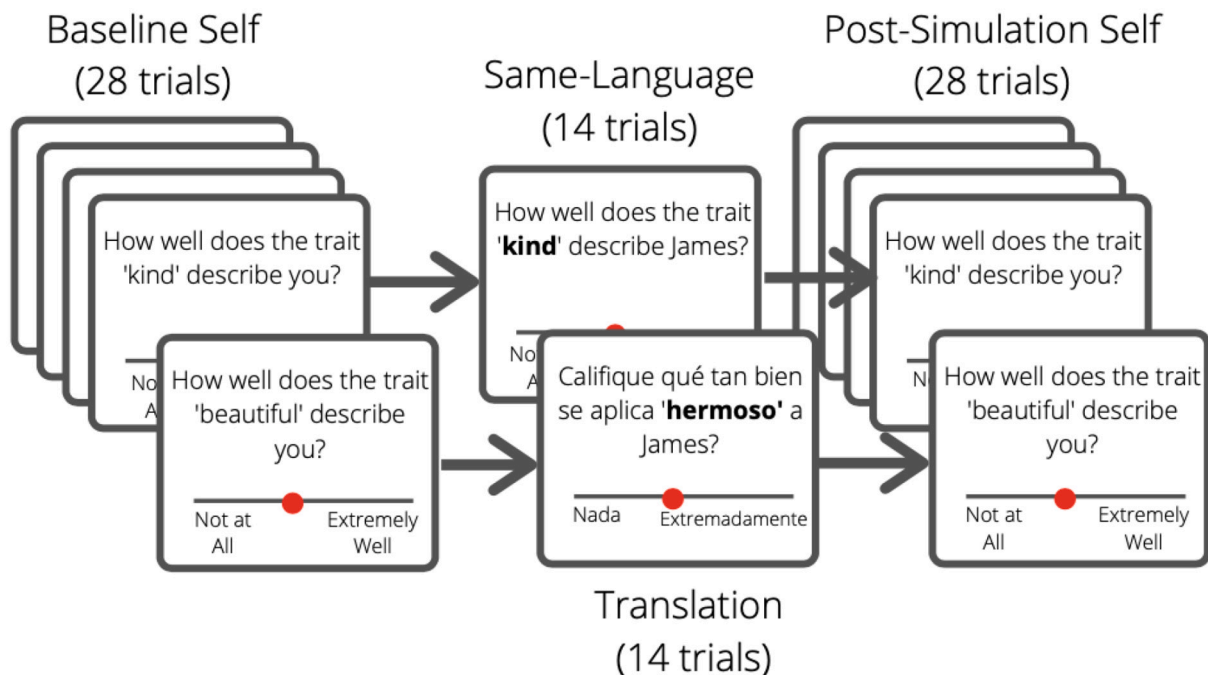
Fig. 9. Task Schematic for Study 5. Participants began by rating themselves on 28 traits, presented in the base language. They then rated a similar other on the same-language trait for half of the original set of traits, and the translation for the other half. The ratings for same-language and translated traits were presented in two blocks, with the order randomized between participants. Participants then re-rated themselves on the 28 initial traits, again in the base language.

highest cosine similarity across the languages were selected for use. For a complete list of trait pairs, please see the Supplement.

If, as in Studies 3–4, simulation activates the semantic concept associated with a trait word, we would expect a main effect of time, across both same-language and translated traits. We would also expect a simple effect of time within both the same-language and the translated traits condition, and no interaction between time and condition.

6.2. Results

We tested for SIM using a 2 (time: baseline and post-simulation) x 2 (type of word simulated: same-language or translated words) repeated-measures ANOVA (Fig. 10). Results support the hypothesis that simulating a target changes self-knowledge at a conceptual level. The ANOVA revealed a significant main effect of time ($F(1,171) = 44.0$, $\eta_p^2 = 0.20$, $p < .001$), such that participants' self-ratings were closer to the target in the post-simulation phase ($M = 17.13$, $SD = 8.08$) than in the baseline phase ($M = 19.37$, $SD = 8.33$). A main effect of type of word was also observed ($F(1,171) = 20.5$, $\eta_p^2 = 0.11$, $p < .001$; same-language $M = 17.09$, $SD = 8.13$; translated $M = 19.41$, $SD = 8.27$). The same-language words showed a smaller difference between the self- and target-ratings. No interaction between these variables was observed ($F(1,171) = 0.66$, $\eta_p^2 = 0.004$, $p = .42$), suggesting that the strength of the effect was similar for both same-language and translated traits.

Additionally, the simple effect of time was calculated within each type of word. We observed significant changes in self-knowledge in the direction of the simulated target for both types of word; the effect size was similar for the same-language ($t(171) = 5.48$, $p < .001$, Cohen's d = 0.26; post-simulation $M = 16.06$, $SD = 7.83$; baseline $M = 18.18$, $SD = 8.32$) and translated words ($t(171) = 5.81$, $p < .001$, Cohen's d = 0.29; post-simulation $M = 18.20$, $SD = 8.21$; baseline $M = 20.61$, $SD = 8.19$).

Together, these results bolster the findings from Studies 3–4 that simulation can change self-knowledge when not only identical, but also semantically related traits are simulated. This finding supports the conclusion that self-knowledge is activated on a conceptual level during simulation of trait knowledge.

7. Discussion

Simulating other people can change how a person thinks about themselves (Meyer et al., 2019). How deeply can simulation change self-knowledge? Six studies provide evidence that simulation induces durable and extensive changes to self-concept. Participants' episodic

Fig. 10. Y-axis represents the difference between self and target ratings; smaller values reflect greater similarity between self and target. Light gray bars reflect similarity before simulation, while dark gray bars reflect similarity after simulation. X-axis represents the type of trait simulated. Error bars represent 95% confidence intervals; * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

memories and trait knowledge were reliably altered to become more like the simulated other. These changes in episodic self-knowledge persisted for at least two days and changes spread through the semantic network of conceptual self-knowledge.

Studies 1 and 2 explored the strength of SIM by studying the duration of the effect. Results support the notion that, within the context of this paradigm, simulation induces durable change in episodic memories. These results provide clarity into how simulation changes the self. If simulation were only capable of changing the working self-concept, any change would have disappeared after a 48-h delay. After such an extended period of time, the working self-concept will have changed; the 'study participant self' will have shifted to the 'mom self' or to the 'coworker self,' as the participants move through different contexts and scenarios over the course of two days. Although participants will have the context of the 'study participant' self both before and after the delay, it is important to note that other contextual factors shape the working self-concept at each time point (e.g., previous situations, social contexts, a person's mood, etc.; Markus & Wurf, 1987). As such, the working self will change in each distinct context, and observing SIM following a delay is important for establishing that social simulation effects change to episodic knowledge on a lasting level, and not just a change to working self-concept.

That said, we observed that the magnitude of the change weakens over time. Thus, the mechanism behind SIM may differ from that of memory reconsolidation effects, which have instead been shown to increase in strength over time (Zhu et al., 2012). Perhaps altering memories about the self requires further reinforcement. That is, the self may be slightly less malleable than other knowledge content such that it needs misinformation to be further reinforced before it embeds itself permanently into one's memories or self-concept. Future research should establish the boundary conditions of SIM by including longer delays, or testing self-knowledge across multiple contexts.

The results from Studies 3 and 4 provide evidence that simulation activates conceptual self-knowledge, thus rendering this self-knowledge – as well as the network of connected knowledge – susceptible to change. Simulating one trait shaped self-ratings on semantically related traits – both synonyms and semantically related traits within one language, and translations across languages. That is, the effect of simulation spreads through the semantic network, rather than remaining confined to the specifically targeted concept. These findings provide evidence that semantic self-knowledge is structured in a manner similar to existing models of semantic knowledge, thus strengthening proposals that semantic self-knowledge is organized in networks arranged by semantic similarity (Elder et al., 2020; Greenwald & Banaji, 1989).

Future research should explore the extent to which SIM can cause lasting changes in other forms of self-knowledge. In Studies 1–2, we observed that simulation has a durable effect on one's episodic self-memories. To what extent does this finding extend to semantic self-knowledge? While episodic memories are prone to change, trait self-knowledge may be less prone to change. That is, the literature on trait-based self-concept suggests that this form of self-knowledge may be more reliably stable (Church et al., 2012; Marsh, Smith, Barnes, & Butler, 1983; McCrae & Costa, 1988). If trait knowledge is indeed more stable than episodic memories, simulation may be capable of changing episodic memories on a longer timescale than it can change semantic self-knowledge. This possibility is supported by findings that trait-related and episodic self-knowledge are distinguishable: in patients with anterograde amnesia, episodic memories were inaccessible, but patients reported accurate trait-based self-ratings, as compared to trait-ratings made following the amnesiac period (Klein et al., 1996). If episodic and semantic self-knowledge are distinguishable to such a degree, they may be differentially activated or changed as a result of simulation. If that is the case, it is possible that simulation causes durable changes to episodic self-knowledge, but only momentary changes to the semantic self-knowledge that comprises working self-concept. Should future studies observe that simulation-induced changes in trait
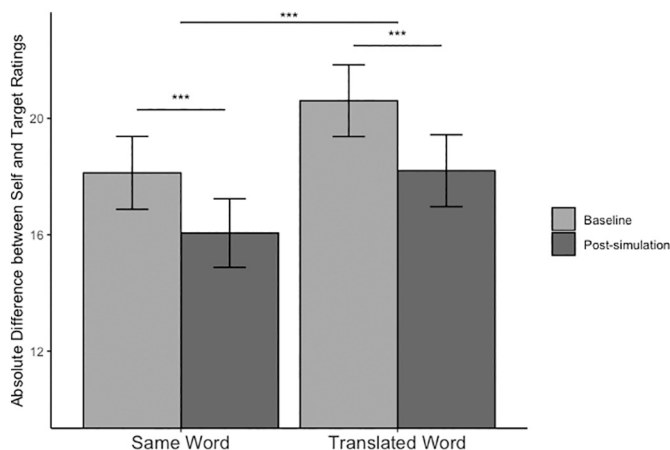
self-knowledge also persist following a delay, this would enhance our understanding of the stability of semantic self-knowledge. Future studies may also consider exploring the extent to which simulation in one aspect of self-knowledge can affect self-knowledge in another area. For example, might simulating another individual on traits change self-knowledge on episodic memories? Such studies may further help us understand the relationship between episodic memories and semantic self-knowledge in self-concept (Conway, 2005; Sakaki, 2007a).

Such results open the door to new possibilities of using social simulation to update the self in a way previously explored only via direct social experiences. For example, we know that direct experience with a role model can benefit underrepresented or disadvantaged students. Women in STEM fields, for instance, benefit from the presence of other women in STEM as role models (Stout, Dasgupta, Hunsinger, & McManus, 2011). In particular, the presence of a role model can lead to increased interest and sense of belonging in STEM (González-Pérez, Mateos de Cabo, & Sáinz, 2020), lower endorsement of gender-STEM stereotypes (Van Camp, Gilbert, & O'Brien, 2019), and higher grades and retention rates in STEM courses (Herrmann et al., 2016). SIM offers an exciting way to explore if these or similar effects can be induced without the presence of other people. Could merely simulating other individuals similarly affect peoples' self-esteem, affect, or feelings of belonging? If simply simulating a role model offers real benefits, then simulation may offer members of underrepresented groups access to some of these benefits.

One unexpected finding in these results was observing similar SIM effects for both similar and dissimilar targets. Prior work suggested that the effects of SIM should be larger when a similar other, rather than a stranger or dissimilar other, is simulated (Meyer et al., 2019). These earlier findings were in line with previous work suggesting that individuals recruit more self-relevant knowledge when considering similar others than dissimilar others (Mashek, Aron, & Boncimino, 2003; Tamir & Mitchell, 2013). The current results showing similar effects for both targets are not unprecedented, however, as several previous studies on SIM also did not observe an interaction (Meyer et al., 2019). We suspect that this finding may have resulted because participants in the current studies included residents of the United States of America, the United Kingdom, and other countries, while previous research used only USA-based participants. In fact, the vast majority of the participants in the studies with multiple targets (70–90%) were not residents of the United States. It may be that given political polarization at the time of data collection (late 2019 and early 2020, prior to the US Presidential Election) and heterogeneity of the sample, non-US based participants may have had a stronger sense of similarity to the target than US-participants, which would skew any interaction. Furthermore, while we assumed that participants across countries would activate the same amount of self-knowledge when simulating the average American, this assumption has not been tested. That said, the amount of self-relevant knowledge activated for each target was not directly manipulated in our studies. Future research should build on this current work to determine the role of activating self-knowledge in SIM. Future work should also look to determine the effectiveness of SIM in various contexts and with more diverse populations, as the participants recruited on the crowdsourcing platform Prolific are not fully representative of the US or international populations.

These findings are also limited in their scope. All of the studies reported here were conducted in the laboratory under highly contrived, controlled contexts. Future work should determine the degree to which this effect is observed outside of laboratory and impacts individuals' lasting attitudes and behaviors. Such research may provide insight into the regularity with which SIM occurs. Simulation, itself, occurs frequently during spontaneous thought (Christoff et al., 2009). It is still an open question, however, under which naturalistic contexts spontaneous simulation results in SIM. Are the effects of SIM good or bad for individuals? One could intentionally leverage SIM to induce positive changes, such as inspiring individuals, increasing their sense of

confidence, or reducing negative memories. However, if SIM occurs frequently, and one is consistently surrounded by individuals with undesirable traits, it is easy to see how SIM could be harmful.

This research begins to clarify how deeply simulation can change self-knowledge, but leaves open the question of *how* it effects this change. For SIM to occur, we suspect that there are multiple component processes that must take place: the initial activation of self-knowledge, the activation of knowledge about the target, and the integration of knowledge about the target with self-knowledge. Future research needs to manipulate each of these components individually to determine which are necessary or sufficient for SIM to occur. For example, is it necessary for individuals to initially activate self-knowledge? We can test this by comparing SIM effects for individuals who initially recall memories prior to simulation, as in the current paradigm, to individuals who do not first activate self-knowledge. If self-activation is necessary, then any adaptation of this paradigm that removes the initial self-rating phase would show no effect of SIM. Is it necessary for the retrieved episodic memory to be recalled in an autobiographical context? We can test this by manipulating the perspective taken by the participant; if a first-person, engaged perspective increases the strength of SIM, then we will know that it is heightened activation of self-knowledge in particular that induces SIM.

It is important to consider the magnitude to which the self-concept shifts as a result of simulation. The changes observed in these studies were measured on a hundred-point rating scale, and self-ratings generally shifted ~2 points on this scale. This finding supports prior work suggesting that the broader self-concept is largely stable (Markus & Wurf, 1987; Oyserman & Elmore, 2012; Shavelson, Hubner, & Stanton, 1976). That is, that individual episodes of simulation do not cause drastic changes in ones' personality or memories. This should be a comfort to those of us who hope for consistency in the self over time. However, we know little about how these small changes may accumulate over time. The majority of people's spontaneous thoughts are social. That is, individuals simulate or think deeply about other individuals often throughout their daily lives (Mar, Mason, & Litvack, 2012; Mildner & Tamir, 2018; Song & Wang, 2012). People simulate others when making decisions, viewing advertisements, or considering counterfactual futures (Baumeister, Hofmann, Summerville, Reiss, & Vohs, 2020; Beaty, Seli, & Schacter, 2019; Elder et al., 2020; Kane et al., 2007). There are thus many opportunities for the effects of simulation may accumulate over time. The direction of the resulting shift will depend on the people and the social context that surround a simulator. Future work should investigate the potential for these kinds of lasting consequences of simulation.

Simulation might induce change not only in the self-concept, but in connections between the self and the simulated target. For instance, as self-concept shifts towards a simulated target, this may serve to foster an increased sense of closeness or similarity. Indeed, prior work on SIM suggested that affiliation goals may drive SIM (Meyer et al., 2019). Future research should consider if the reverse may all be true, that simulation drives enhanced interpersonal affiliation. Prior research on simulation suggests that this may be the case; when simulations are more detailed, empathy for both ingroup and outgroup members increases (Vollberg, Gaesser, & Cikara, 2021). Such potential downstream changes in social relationships and representations of others should be considered when exploring both the mechanisms behind and potential applications of SIM.

Our thoughts reflect the external world around us. Humans are social creatures who think about other people even when alone. The people around us in our real social world are reflected in the content of our thought (Mildner & Tamir, 2018). Here we find that, by thinking about those other people, individuals may be inadvertently updating their sense of self to reflect the other person. Moreover, we find that these changes to self-concept can be lasting and expansive. These findings suggest that the company we keep is vital for shaping not only our direct interactions with the social world, but for determining our very sense of

self.

Open practices

Materials and data for all experiments are available on OSF (https://osf.io/ga65b/). Pre-registrations are available on OSF for Study 3 (https://osf.io/t8bg2), Study 4 (https://osf.io/ah8gz), and Study 5 (https://osf.io/dpafe).

Author note

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2021.104229.

References

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review, 5*(1), 1–21.

Baumeister, R. F., Hofmann, W., Summerville, A., Reiss, P. T., & Vohs, K. D. (2020). Everyday thoughts in time: Experience sampling studies of mental time travel. *Personality and Social Psychology Bulletin, 46*(12), 1631–1648. https://doi.org/10.1177/0146167220908411.

Beaty, R. E., Seli, P., & Schacter, D. L. (2019). Thinking about the past and future in daily life: An experience sampling study of individual differences in mental time travel. *Psychological Research, 83*(4), 805–816. https://doi.org/10.1007/s00426-018-1075-7.

Bloom, K. L. (1961). Age and the self concept. *American Journal of Psychiatry, 118*(6), 534–538. https://doi.org/10.1176/ajp.118.6.534.

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics, 5*, 135–146, 1607.04606 [Cs] http://arxiv.org/abs/1607.04606.

Boomsma, C., Pahl, S., & Andrade, J. (2016). Imagining change: An integrative approach toward explaining the motivational role of mental imagery in pro-environmental behavior. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01780.

Borghuis, J., Denissen, J. J. A., Sijtsma, K., Branje, S., Meeus, W. H. J., & Bleidorn, W. (2018). Positive daily experiences are associated with personality trait changes in middle-aged mothers: Daily experiences and personality change. *European Journal of Personality, 32*(6), 672–689. https://doi.org/10.1002/per.2178.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences, 11*(2), 49–57. https://doi.org/10.1016/j.tics.2006.11.004.

Calkins, M. W. (1908). Psychology as science of self: I. is the self body or has it body? *The Journal of Philosophy, Psychology and Scientific Methods, 5*(1), 12. https://doi.org/10.2307/2012278.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences, 106*(21), 8719–8724. https://doi.org/10.1073/pnas.0900234106.

Church, T. A., Alvarez, J. M., Katigbak, M. S., Mastor, K. A., Cabrera, H. F., Tanaka-Matsumi, J., … Buchanan, A. L. (2012). Self-concept consistency and short-term stability in eight cultures. *Journal of Research in Personality, 46*(5), 556–570. https://doi.org/10.1016/j.jrp.2012.06.003.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407.

Conway, M. A. (2005). Memory and the self☆. *Journal of Memory and Language, 53*(4), 594–628. https://doi.org/10.1016/j.jml.2005.08.005.

Conway, M. A., & Bekerian, D. A. (1987). Organization in autobiographical memory. *Memory & Cognition, 15*(2), 119–132. https://doi.org/10.3758/BF03197023.

Conway, M. A., Justice, L. V., D'Argembeau, A., Conway, M. A., Justice, L. V., & D'Argembeau, A. (2019). The self-memory system revisited: Past, present, and future. In *The organization and structure of autobiographical memory* (pp. 28–51). Oxford University Press. https://doi.org/10.1093/oso/9780198784845.003.0003.

Cooley, C. (1902). *Human nature and the social order* (C. Scribner's Sons).

Davies, M. (2008). The Corpus of Contemporary American English (COCA): One billion words, 1990-2019. https://www.english-corpora.org/coca/.

Davies, M. (2016). Corpus del Español: Two billion words, 21 countries. http://www.corpusdelespanol.org/web-dial/.

Demo, D. H. (1992). The self-concept over time: Research issues and directions. *Annual Review of Sociology, 18*(1), 303–326. https://doi.org/10.1146/annurev.so.18.080192.001511.

DeSteno, D. A., & Salovey, P. (1997). The effects of mood on the structure of the self-concept. *Cognition & Emotion, 11*(4), 351–372. https://doi.org/10.1080/026999397379836.

Di Bella, L., & Crisp, R. J. (2015). Imagining oneself in a stereotyped role may stifle generalized tendencies to support social change. *Social Influence, 10*(3), 157–167. https://doi.org/10.1080/15534510.2015.1030446.

Diehl, M., Jacobs, L. M., & Hastings, C. T. (2006). Temporal stability and authenticity of self-representations in adulthood. *Journal of Adult Development, 13*(1), 10–22. https://doi.org/10.1007/s10804-006-9001-4.

Downey, L., Rosengren, D. B., & Donovan, D. M. (2000). To thine own self be true. *Addictive Behaviors, 25*(5), 743–757. https://doi.org/10.1016/S0306-4603(00)00091-5.

Dumas, J. E., Johnson, M., & Lynch, A. M. (2002). Likableness, familiarity, and frequency of 844 person-descriptive words. *Personality and Individual Differences, 32*(3), 523–531. https://doi.org/10.1016/S0191-8869(01)00054-X.

Elder, J., Cheung, K. M. B. M. B., Davis, T., & Hughes, B. (2020). Mapping the Self: A network approach for understanding psychological and neural representations of self-concept structure [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/hj87w.

Ellemers, N., Kortekaas, P., & Ouwerkerk, J. W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology, 29*(23).

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General, 120*(4), 339–357. https://doi.org/10.1037/0096-3445.120.4.339.

Fazio, R. H., Effrein, E. A., & Falender, V. J. (1981). Self-perceptions following social interaction. *Journal of Personality and Social Psychology, 41*(2), 232–242. https://doi.org/10.1037/0022-3514.41.2.232.

Fleeson, W. (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality, 75*(4), 825–862. https://doi.org/10.1111/j.1467-6494.2007.00458.x.

Frenck, C., & Pynte, J. (1987). Semantic representation and surface forms: A look at across-language priming in bilinguals. *Journal of Psycholinguistic Research, 16*(4), 383–396. https://doi.org/10.1007/BF01069290.

Gaesser, B., & Schacter, D. L. (2014). Episodic simulation and episodic memory can increase intentions to help others. *Proceedings of the National Academy of Sciences, 111*(12), 4415–4420. https://doi.org/10.1073/pnas.1402461111.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493–501. https://doi.org/10.1016/S1364-6613(98)01262-5.

Gilead, M., Boccagno, C., Silverman, M., Hassin, R. R., Weber, J., & Ochsner, K. N. (2016). Self-regulation via neural simulation. *Proceedings of the National Academy of Sciences, 113*(36), 10037–10042. https://doi.org/10.1073/pnas.1600159113.

González-Pérez, S., Mateos de Cabo, R., & Sáinz, M. (2020). Girls in STEM: Is it a female role-model thing? *Frontiers in Psychology, 11*, 2204. https://doi.org/10.3389/fpsyg.2020.02204.

Gore, J. S. (2005). *Determinants of self-concept change in new environments.* Iowa State University.

Gray, K., Anderson, S., Chen, E. E., Kelly, J. M., Christian, M. S., Patrick, J., … Lewis, K. (2019). "Forward flow": A new measure to quantify free thought and predict creativity. *American Psychologist, 74*(5), 539–554. https://doi.org/10.1037/amp0000391.

Greenwald, A. G., & Banaji, M. R. (1989). The self as a memory system: Powerful, but ordinary. *Journal of Personality and Social Psychology, 57*(1), 41–54. https://doi.org/10.1037/0022-3514.57.1.41.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*(3), 225–240. https://doi.org/10.1037/0278-7393.6.3.225.

Herrmann, S. D., Adelman, R. M., Bodford, J. E., Graudejus, O., Okun, M. A., & Kwan, V. S. Y. (2016). The effects of a female role model on academic performance and persistence of women in STEM courses. *Basic and Applied Social Psychology, 38* (5), 258–268. https://doi.org/10.1080/01973533.2016.1209757.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review, 119*(2), 431–440. https://doi.org/10.1037/a0027373.

Hinkley, K., & Andersen, S. M. (1996). The working self-concept in transference: Significant-other activation and self change. *Journal of Personality and Social Psychology, 71*(6), 1279–1295. https://doi.org/10.1037/0022-3514.71.6.1279.

Joulin, A., Bojanowski, P., Mikolov, T., Jegou, H., & Grave, E. (2018). Loss in translation: Learning bilingual word mapping with a retrieval criterion. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing,*, 1804.07745 [Cs] http://arxiv.org/abs/1804.07745.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18*(7), 614–621. https://doi.org/10.1111/j.1467-9280.2007.01948.x.

Klein, S. B., Chan, R. L., & Loftus, J. (1999). Independence of episodic and semantic self-knowledge: The case from autism. *Social Cognition, 17*(4), 413–436. https://doi.org/10.1521/soco.1999.17.4.413.

Klein, S. B., & Lax, M. L. (2010). The unanticipated resilience of trait self-knowledge in the face of neural damage. *Memory, 18*(8), 918–948. https://doi.org/10.1080/09658211.2010.524651.

Klein, S. B., Loftus, J., & Kihlstrom, J. F. (1996). Self-knowledge of an amnesic patient: Toward a neuropsychology of personality and social psychology. *Journal of Experimental Psychology: General, 125*(3), 250–260. https://doi.org/10.1037/0096-3445.125.3.250.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008.

Libby, L. K., Valenti, G., Pfent, A., & Eibach, R. P. (2011). Seeing failure in your life: Imagery perspective determines whether self-esteem shapes reactions to recalled and imagined failure. *Journal of Personality and Social Psychology, 101*(6), 1157–1173. https://doi.org/10.1037/a0026105.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361–366. https://doi.org/10.1101/lm.94705.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(1), 19–31. https://doi.org/10.1037/0278-7393.4.1.19.

Lommen, M. J. J., Engelhard, I. M., & van den Hout, M. A. (2013). Susceptibility to long-term misinformation effect outside of the laboratory. *European Journal of Psychotraumatology, 4*(1), 19864. https://doi.org/10.3402/ejpt.v4i0.19864.

Mar, R. A., Mason, M. F., & Litvack, A. (2012). How daydreaming relates to life satisfaction, loneliness, and social support: The importance of gender and daydream content. *Consciousness and Cognition, 21*(1), 401–407. https://doi.org/10.1016/j.concog.2011.08.001.

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology, 29*(1), 87–109. https://doi.org/10.1006/jesp.1993.1005.

Markman, K. D., Klein, W. M. P., & Suhr, J. A. (2012). *Handbook of imagination and mental simulation* (0 ed.). Psychology Press. https://doi.org/10.4324/9780203809846.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38*(1), 299–337. https://doi.org/10.1146/annurev.ps.38.020187.001503.

Marsh, H. W., Smith, I. D., Barnes, J., & Butler, S. (1983). Self-concept: Reliability, stability, dimensionality, validity, and the measurement of change. *Journal of Educational Psychology, 75*(5), 772–790. https://doi.org/10.1037/0022-0663.75.5.772.

Mashek, D. J., Aron, A., & Boncimino, M. (2003). Confusions of self with close others. *Personality and Social Psychology Bulletin, 29*(3), 382–392. https://doi.org/10.1177/0146167202250220.

McCrae, R. R., & Costa, P. T. (1988). Age, personality, and the spontaneous self-concept. *Journal of Gerontology, 43*(6), S177–S185. https://doi.org/10.1093/geronj/43.6.S177.

McRae, K. (2004). Semantic memory: Some insights from feature-based connectionist attractor networks. In , *45. Psychology of learning and motivation* (pp. 41–86). Elsevier. https://doi.org/10.1016/S0079-7421(03)45002-0.

Meevissen, Y. M. C., Peters, M. L., & Alberts, H. J. E. M. (2011). Become more optimistic by imagining a best possible self: Effects of a two week intervention. *Journal of Behavior Therapy and Experimental Psychiatry, 42*(3), 371–378. https://doi.org/10.1016/j.jbtep.2011.02.012.

Meyer, M. L., Zhao, Z., & Tamir, D. I. (2019). Simulating other people changes the self. *Journal of Experimental Psychology: General, 148*(11), 1898–1913. https://doi.org/10.1037/xge0000565.

Mildner, J. N., & Tamir, D. (2018). *The people around you are inside your head: Social context shapes spontaneous thought [Preprint]*. PsyArXiv. https://doi.org/10.31234/osf.io/xmzh7.

Mildner, J. N., & Tamir, D. I. (2019). Spontaneous thought as an unconstrained memory process. *Trends in Neurosciences, 42*(11), 763–777. https://doi.org/10.1016/j.tins.2019.09.001.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM, 38*(11), 39–41. https://doi.org/10.1145/219717.219748.

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 364*(1521), 1309–1316. https://doi.org/10.1098/rstb.2008.0318.

Nickerson, R. S. (1999). How we know—And sometimes misjudge—What others know: Imputing one's own knowledge to others. *Psychological Bulletin, 125*(6), 737–759. https://doi.org/10.1037/0033-2909.125.6.737.

Ost, J., Foster, S., Costall, A., & Bull, R. (2005). False reports of childhood events in appropriate interviews. *Memory, 13*(7), 700–710. https://doi.org/10.1080/09658210444000340.

Oyserman, D., & Elmore, K. (2012). Self, self-concept, and identity. In *Handbook of self and identity* (2nd ed., pp. 69–104).

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition, 62*(2), 223–240. https://doi.org/10.1016/S0010-0277(96)00782-2.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research, 66*(3), 180–194. https://doi.org/10.1007/s00426-002-0086-5.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116*(1), 129–156. https://doi.org/10.1037/a0014420.

Purkey, W. (1988). *An overview of self-concept theory for counselors*. Highlights: An ERIC/CAPS Digest.

Saint-Aubin, J., Ouellette, D., & Poirier, M. (2005). Semantic similarity and immediate serial recall: Is there an effect on all trials. *Psychonomic Bulletin & Review, 12*(1), 171–177. https://doi.org/10.3758/BF03196364.

Sakaki, M. (2007a). Semantic self-knowledge and episodic self-knowledge: Independent or interrelated representations? *Memory, 15*(1), 1–16. https://doi.org/10.1080/09658210601055750.

Sakaki, M. (2007b). Mood and recall of autobiographical memory: The effect of focus of self-knowledge. *Journal of Personality, 75*(3), 421–450. https://doi.org/10.1111/j.1467-6494.2007.00444.x.

Schacter, D. L., & Addis, D. R. (2007). On the constructive episodic simulation of past and future events. *Behavioral and Brain Sciences, 30*(3), 331–332. https://doi.org/10.1017/S0140525X07002178.

Schmolck, H., Buffalo, E. A., & Squire, L. R. (2000). Memory distortions develop over time: Recollections of the O.J. Simpson trial verdict after 15 and 32 months. *Psychological Science, 11*(1), 39–45. https://doi.org/10.1111/1467-9280.00212.

Segal, Z. V., & Vella, D. D. (1990). Self-schema in major depression: Replication and extension of a priming methodology. *Cognitive Therapy and Research, 14*(2), 161–176. https://doi.org/10.1007/BF01176207.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407–441. https://doi.org/10.3102/00346543046003407.

Song, X., & Wang, X. (2012). Mind wandering in Chinese daily lives – An experience sampling study. *PLoS One, 7*(9), Article e44423. https://doi.org/10.1371/journal.pone.0044423.

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology, 100*(2), 255–270. https://doi.org/10.1037/a0021385.

Swann, W. B., & Hill, C. A. (1982). When our identities are mistaken: Reaffirming self-conceptions through social interaction. *Journal of Personality and Social Psychology, 43*(1), 59–66. https://doi.org/10.1037/0022-3514.43.1.59.

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General, 142*(1), 151–162. https://doi.org/10.1037/a0028232.

Taylor, S. E., Pham, L. B., Rivkin, I. D., & Armor, D. A. (1998). Harnessing the imagination: Mental simulation, self-regulation, and coping. *American Psychologist, 53*(4), 429–439. https://doi.org/10.1037/0003-066X.53.4.429.

Thomas, A. K., Hannula, D. E., & Loftus, E. F. (2007). How self-relevant imagination affects memory for behaviour. *Applied Cognitive Psychology, 21*(1), 69–86. https://doi.org/10.1002/acp.1270.

Trafimow, D., Silverman, E. S., Fan, R. M.-T., & Fun Law, J. S. (1997). The effects of language and priming on the relative accessibility of the private self and the collective self. *Journal of Cross-Cultural Psychology, 28*(1), 107–123. https://doi.org/10.1177/0022022197281007.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138–146. https://doi.org/10.1037/0894-4105.11.1.138.

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of self-esteem across the life span. *Journal of Personality and Social Psychology, 84*(1), 205–220.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114.

Van Camp, A. R., Gilbert, P. N., & O'Brien, L. T. (2019). Testing the effects of a role model intervention on women's STEM outcomes. *Social Psychology of Education, 22*(3), 649–671. https://doi.org/10.1007/s11218-019-09498-2.

Vollberg, M. C., Gaesser, B., & Cikara, M. (2021). Activating episodic simulation increases affective empathy. *Cognition, 209*, 104558. https://doi.org/10.1016/j.cognition.2020.104558.

Wilson, R. E., Thompson, R. J., & Vazire, S. (2017). Are fluctuations in personality states more than fluctuations in affect? *Journal of Research in Personality, 69*, 110–123. https://doi.org/10.1016/j.jrp.2016.06.006.

Zhu, B., Chen, C., Loftus, E. F., He, Q., Chen, C., Lei, X., … Dong, Q. (2012). Brief exposure to misinformation can Lead to long-term false memories: Long-term false memory. *Applied Cognitive Psychology, 26*(2), 301–307. https://doi.org/10.1002/acp.1825.