Mechanisms Underlying the Prosocial Construction Effect: Detail and Self-Referential Processing

by

Caspian Sawczak

A thesis submitted in conformity with the requirements for the degree of Master of Arts
Psychology
University of Toronto

© Copyright by Caspian Sawczak 2015
Mechanisms Underlying the Prosocial Construction Effect: Detail and Self-Referential Processing

Caspian Sawczak

Master of Arts

Psychology

University of Toronto

2015

Abstract

There is growing interest in the links between memory and imagination and in their contributions to other cognitive functions. In previous research, imagining oneself helping others increased one’s willingness to help them, and willingness to help correlated with the subjective vividness of imagination (i.e., the prosocial construction effect; PCE). We investigated the degree to which the PCE relies on self-referential processing and on more objectively measured vividness, and the relationship between episodic memory and self-reported trait empathy. We conclude that 1) self-reference is an important but not a necessary component of the PCE; 2) its role may relate to increased feelings of compassion; 3) the PCE is not limited to subjective vividness but can be demonstrated using more objective measures of vividness; and 4) self-reported trait empathy is higher when survey items cue memories of one’s own empathetic behaviour. Implications and avenues for future research are discussed.
Table of Contents

Contents

Table of Contents ........................................................................................................................................ iii
List of Tables ............................................................................................................................................... vi
List of Figures ............................................................................................................................................. vii
List of Appendices ................................................................................................................................... viii
1 Introduction ............................................................................................................................................. 1
  1.1 The Prosocial Construction Effect (PCE) ......................................................................................... 1
  1.2 Questions Warranting Further Attention ......................................................................................... 3
    1.2.1 Does episodic detail play a critical role distinct from that of semantic detail? ..... 3
    1.2.2 Would excluding oneself from the episodic representation diminish the effect? ... 5
  1.3 Overview of the Present Study (Experiments 1 and 2) ............................................................... 7
2 Method: Experiment 1 ............................................................................................................................. 8
  2.1 Participants ......................................................................................................................................... 8
  2.2 Materials and Design ....................................................................................................................... 9
    2.2.1 Stimuli ......................................................................................................................................... 9
    2.2.2 Design ....................................................................................................................................... 10
    2.2.3 Instruments .............................................................................................................................. 10
  2.3 Procedure .......................................................................................................................................... 11
    2.3.1 Instruction phase ...................................................................................................................... 11
    2.3.2 Task phase .............................................................................................................................. 13
    2.3.3 Question phase ....................................................................................................................... 14
3 Results: Experiment 1 ............................................................................................................................ 15
  3.1 Willingness to help ........................................................................................................................... 15
  3.2 Subjective vividness ratings ............................................................................................................. 16
3.3 Relationship between vividness and willingness to help ................................................. 17
3.4 Confidence and emotion ratings ..................................................................................... 18
4 Discussion: Experiment 1 ...................................................................................................... 19
5 Method: Experiment 2 ........................................................................................................... 22
  5.1 Participants ......................................................................................................................... 22
  5.2 Materials and Design ......................................................................................................... 22
    5.2.1 Stimuli ............................................................................................................................ 22
    5.2.2 Design ........................................................................................................................... 22
    5.2.3 Toronto Empathy Questionnaire (TEQ) ........................................................................ 23
  5.3 Procedure ............................................................................................................................. 24
    5.3.1 Instruction phase .......................................................................................................... 24
    5.3.2 Task phase .................................................................................................................... 25
    5.3.3 Question phase ............................................................................................................. 25
  5.4 Scoring Transcripts ............................................................................................................. 26
6 Results: Experiment 2 ........................................................................................................... 27
  6.1 Willingness to help ............................................................................................................. 27
  6.2 Subjective vividness ratings ............................................................................................... 29
  6.3 Transcript scores (episodicity) .......................................................................................... 29
  6.4 Relationships among vividness ratings, internal detail, and willingness to help .............. 31
    6.4.1 Subjective vividness ratings ......................................................................................... 31
    6.4.2 Internal detail content in transcripts ............................................................................ 32
7 Discussion: Experiment 2 ...................................................................................................... 33
  7.1 On subjective vividness ratings ......................................................................................... 34
  7.2 On transcripts and internal details ................................................................................... 36
  7.3 On self-reported trait empathy (TEQ scores) ................................................................... 39
8 Introduction: Experiment 3 ................................................................................................... 40
List of Tables

Table 1. Means for subjective ratings associated with imagined helping scenarios, by task type, in Experiment 1 ................................................................. 17

Table 2. Correlation matrix comparing subjective ratings of vividness to internal detail scores in transcripts of imagined helping scenarios in Experiment 2 ........................................... 33

Table 3. Comparative summary of results from Experiments 1 and 2 ................................................................. 44
List of Figures

Figure 1. Mean willingness to help by task type in Experiment 1................................................16

Figure 2. Mean subjective vividness ratings plotted against mean willingness to help............... in all three conditions of Experiment 1, at the participant level...............................18

Figure 3. Mean willingness to help by task type in Experiment 2...............................................28

Figure 4. Mean willingness to help plotted against TEQ score, by task type, in.............................Experiment 2.................................................................28

Figure 5. Mean internal details as proportions of total details (total number of......................... internal + total number of external) by task type in Experiment 2........................................30

Figure 6. Mean subjective vividness ratings plotted against mean willingness to help................. in both conditions of Experiment 2, at the participant level.................................31
List of Appendices

Appendix A – Examples of Vignettes and Math Problems.................................................................57

Appendix B – Memory Follow-Up Survey (MemFUS), page 1.............................................................58
1 Introduction

Humans have a remarkable ability to mentally reconstruct and re-experience specific past events. Memories that are recollected in this way are known in the scientific literature as episodic memories and are typically contrasted with semantic memories, or memories for facts and general knowledge not tied to any particular spatial or temporal context (Tulving, 1972, 2002). There is ample evidence that episodic memory recollection is supported by a network of neural structures including the hippocampus and other parts of the medial temporal lobe (MTL; for reviews, see Maguire, 2001; Svoboda, McKinnon, & Levine, 2006). It is also well established that recollecting events relies on some of the same cognitive processes and brain areas (including the MTL) as does imagining events (for review, see Schacter et al., 2012 or Mullally & Maguire, 2014), which has led to discussion of the general capacity to mentally construct scenes and represent events in a detailed, coherent, narrative (e.g., Hassabis & Maguire, 2007; Rubin & Umanath, 2014), which we will refer to here as episodic representation.

There is a growing body of evidence that episodic representation is an important component of other cognitive functions, such as making decisions regarding delayed versus immediate rewards (Benoit, Gilbert, & Burgess, 2011), solving open-ended problems (Sheldon, McAndrews, & Moscovitch, 2011; Sheldon, Vandermorris, Al-Haj, Cohen, Winocur, & Moscovitch, 2015), displaying empathy (Ciaramelli, Bernardi, & Moscovitch, 2013), and fostering intentions to help other people (Gaesser & Schacter, 2014). The present study (Experiments 1 and 2) extends work that has been conducted on the latter phenomenon. First we will describe what has been demonstrated by previous research. Then we will discuss some of the questions that warrant further attention, reviewing relevant literature in the process. This will be followed by a brief summary of the literature review and a description of the first two experiments we conducted. Later on in the paper (i.e., after presenting Experiments 1 and 2), we will introduce a third experiment that address the related issue of self-reported trait empathy.

1.1 The Prosocial Construction Effect (PCE)

In a recent study, participants reported being more willing to help individuals in hypothetical situations of need after either vividly imagining themselves helping those individuals or recalling past experiences in which they helped others in related situations, compared to various control
conditions (Gaesser & Schacter, 2014). As this effect (which we have termed and will refer to hereafter as the *prosocial construction effect*) was demonstrated in three very similar experiments, the following overview will collapse across them.

First, participants were shown a series of one-sentence vignettes, one at a time, each presenting a nondescript, anonymous individual faced with some kind of problem (e.g., “Driving to their wedding, this person’s car broke down on the highway”; “This person's dog has not returned home in the last 24 hours”). After reading a vignette, participants would spend one minute either 1) vividly imagining themselves helping the person in need; 2) coming up with specific strategies for helping the person in need and imagining them being described in text on a web page but *not* imagining the strategies themselves being carried out; 3) recalling a past personal experience in which they helped someone in a similar situation; 4) analyzing the journalistic writing style of the vignette; or 5) solving mathematical word problems completely unrelated to the vignette. Next, participants were shown each vignette again, only this time they rated (on a scale from 1 to 7) how willing they were to help the person in need (i.e., prosocial intent); the degree to which they experienced various emotions in response to the story; the degree to which they considered the thoughts and feelings of the person in the story (i.e., perspective-taking); and the vividness of the events they imagined or recalled (depending on which condition the vignette had been assigned to in the first part of the experiment). They also provided brief descriptions of what they imagined or recalled, where appropriate, and, in the case of remembered events, also rated the similarity of the past experience to the situation described in the present vignette. Results indicated that vividly imagining or recalling instances of helping someone (i.e., episodic representation) led to higher ratings of prosocial intent compared to the three control conditions. There was no statistically significant difference between imagining and remembering, nor were there any significant differences among control conditions. Interestingly, participants’ ratings of the vividness of their imagined or remembered events predicted their prosocial intent on a trial-by-trial basis. Finally, there were no consistent relationships between self-reported emotional reactions, nor perspective-taking, and prosocial intent (Gaesser & Schacter, 2014). Thus, it appeared to be the effects of vivid recollection or simulation *per se*, and not evoked emotional or social content, that mediated prosocial intent.
1.2 Questions Warranting Further Attention

1.2.1 Does episodic detail play a critical role distinct from that of semantic detail?

In the study reviewed above, self-reported vividness of imagined and remembered events was operationalized as both sensory detail and coherence; participants rated both qualities for each episode they generated. For simplicity’s sake, here we will focus on the detail measure, defined as the total richness of imagined sensory input. Analyses from multiple experiments in the study reviewed above revealed strong positive correlations between willingness to help and sensory detail in the imagining and remembering conditions (with $r$’s ranging from .43 to .67; all $p$’s < .05; Gaesser & Schacter, 2014). Also, it is clear from visual inspection of scatterplots that episodes that were rated as lacking detail were associated with ratings of willingness to help that were just as low as ratings of willingness to help in conditions where no episodes were imagined or recalled. Perhaps the prosocial construction effect (PCE) hinges on how detailed the episodic representation is. Setting aside, for now, the fact that we cannot conclude from correlational data whether detail plays such a causal role, we should first consider whether we have reason to believe it might.

Past research has shown that the more frequently people imagine performing a particular behaviour, the greater their intentions are to perform that behaviour in the future (Anderson, 1983). Repeatedly imagining future events also leads people to perceive those events as being more plausible (Carroll, 1978; Szpunar & Schacter, 2013). This is consistent with classic cognitive theories on the availability heuristic (Tversky & Kahneman, 1973) and the related simulation heuristic (Kahneman & Tversky, 1982), which state that people make judgements about situations based on how readily relevant past scenarios come to mind, or how easy it is to imagine a scenario, respectively. In the social psychology literature, dozens of studies have been conducted on the imagined contact effect, in which individuals are more willing to interact with members of a social outgroup, and express more positive attitudes towards them, after imagining themselves interacting with those people (for a meta-analytic review, see Miles & Crisp, 2014). Interestingly, this effect is stronger when participants are instructed to imagine those events in greater detail (Husnu & Crisp, 2010; Husnu & Crisp, 2011). Given that people are more likely to recall verbal material if they encode it in more elaborate semantic detail (Craik & Tulving, 1975), it seems possible that the effect of imagining an event in detail would be to create a strong
memory for that imagined event, which would make it more accessible and thus more likely to influence decision-making as per the availability heuristic (Tversky & Kahneman, 1973).

It is important, however, to distinguish between semantic details, which are not specific to any particular imagined or remembered episode, and perceptual or episodic details, which are. Although semantic knowledge may provide a sort of “scaffolding” that is required to support episodic details (Irish, Addis, Hodges, & Pugnet, 2012), there is evidence that older adults (Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002), as well as individuals with epilepsy (St-Laurent, Moscovitch, Levine, & McAndrews, 2009), post-traumatic stress disorder (Brown et al., 2014), and major depression (Söderlund et al., 2014), are impaired in their ability to generate episodic, but not semantic, details when recollecting events. Such deficits in episodic detail typically arise from damage to or dysfunction in the MTL, particularly the hippocampus (e.g., Hassabis, Kumaran, Vann, & Maguire, 2007; Race, Keane, & Verfaellie, 2011), and in healthy people, the level of detail in episodic representations is closely related to hippocampal activity (Addis, Moscovitch, Crawley, & McAndrews, 2004; Addis & Schacter, 2008), suggesting that the hippocampus is important for integrating disparate episodic details into a single coherent episode (Addis, Wong, & Schacter, 2007). In contrast, the retrieval of semantic detail appears to rely on the anterior temporal lobes (e.g., Addis, McIntosh, Moscovitch, Crawley, & McAndrews, 2004; Brambati, Benoit, Monetta, Belleville, & Joubert, 2010). Also, there have been many neuropsychological cases in which either episodic memory is impaired while semantic memory is intact, or vice versa (e.g., Graham, Simons, Pratt, Patterson, & Hodges, 2000; Nestor, Fryer, & Hodges, 2006; Rosenbaum et al., 2005; Smith & Lah, 2011; Temple & Richardson, 2004). Finally, performance on memory tasks can easily be altered by manipulating whether semantic or perceptual cues are presented (e.g., Challis et al., 1993).

Although it is important not to confuse semantic with perceptual detail, it has long been known that perceptual salience enhances recall (e.g., Tayor & Fiske, 1978 as cited in Lord, 1980), and encouraging people to elaborate on the perceptual details of counterfactual events increases their likelihood to falsely recall those events as having actually occurred (Drivdahl & Zaragoza, 2001). Thus, it is reasonable to expect that an imagined event high in perceptual detail, even if containing relatively few semantic details, will create a strong memory trace that can influence judgments in accordance with the availability heuristic (Tversky & Kahneman, 1973).
In light of this, the present study was designed to be sensitive to the contribution of perceptual (and other episodic) details to the PCE distinct and dissociated from that of semantic details. To this end, we will administer the Autobiographical Interview (AI; Levine et al., 2002). The AI was originally developed as a standardized method for distinguishing and assessing the quantity and quality of both episodic and semantic details in remembered events, but it has also been adapted successfully to the analysis of imagined events (e.g., Addis, Wong, & Schacter, 2008; Brown et al., 2014). In the AI, participants are asked to describe an imagined event in as much detail as possible. Their responses are transcribed and analyzed by segmenting the verbal transcript into chunks and classifying each chunk as a certain type of detail.

In addition to distinguishing between episodic and semantic detail, the AI allows for more thorough and objective analysis of the vividness of imagined events compared to what can be achieved through the use of self-report rating scales. That being said, we expect that AI-derived measures of “episodicity” will correlate with subjective ratings of vividness, based not only on intuition but also on unpublished data confirming this relationship (cited in Gilboa, Winocur, Grady, Hevenor, & Moscovitch, 2004).

In short, the present study will modify the paradigm used by Gaesser & Schacter’s (2014) original PCE study, employing an adapted version of the AI in order to measure the vividness of imagined helping scenarios more objectively and to dissociate the contribution of episodic detail to the PCE from that of semantic detail. We will now turn to the second question that the present study will address.

1.2.2 Would excluding oneself from the episodic representation diminish the effect?

In Gaesser & Schacter’s (2014) experiments, control conditions were intended to address potential confounds contributing to social cognition including mental imagery, retrieving information from semantic memory, thinking of examples of helping people in need, and paying extended attention to the textual content of the stories. However, the two episodic representation conditions that produced an effect—imagining oneself helping a person in need, and remembering oneself helping a person in need—both involved self-referential processing, while none of the control conditions did. It is important to examine the role of self-referential processing in this effect because it may shed light on the underlying psychological mechanism.
As discussed in the previous section, the PCE may be attributable to the availability heuristic (Tversky & Kahneman, 1973). This possibility was noted by Gaesser & Schacter (2014) and has also been referred to by others studying the effects of episodic representation on intergroup attitudes (e.g., Husnu & Crisp, 2010). Consider that the availability heuristic relies on effective memory retrieval: the basic idea is that people decide in favour of things that they have previously learned or experienced and that are readily brought to mind (or cued) by the conscious decision-making process, largely as a function of their representativeness or probability (Tversky & Kahneman, 1973). In other words, the decision options act as retrieval cues. As such, their operations are probably subject to factors that are known to influence retrieval, such as the method or “depth” in which the target was encoded (e.g., Craik & Tulving, 1975; Challis, Velichkovsky, & Craik, 1996). Many experiments have shown that material encoded in reference to oneself is particularly easy to recall. A typical example: if a person is shown the word, “courageous”, and is then asked, “Does this word describe you?”, he or she will be more likely to recall that word than if he or she had been asked, “Does this word have a similar meaning to that of ‘brave’?”. This phenomenon is known as the self-reference effect (SRE; Rogers, Kuiper, & Kirker, 1977; for review, see Symons & Johnson, 1997; and Klein, 2012). With regards to the PCE, imagining oneself helping can be expected to create a memory trace (for the imagined event) that is self-referentially encoded and, thus, more likely to be recalled, and thereby to influence one’s decision to help, than would be a comparable memory trace that was not self-referentially encoded, say, the memory of an imagined event in which someone else is helping the person in need. It is possible, then, that the PCE is due, at least in part, to the mnemonic benefits of self-referential encoding.

It should be noted that some studies have reported an absence or even a reversal of the SRE when using mental imagery during encoding (e.g., Lord, 1980, 1987; Maki & McCaul, 1985; Aron, Aron, Tudor, & Nelson, 1991). As Klein (2012) points out, however, these experiments involved mental imagery but not autobiographical memory retrieval (cf. Brown, Keenan, & Potts, 1986, Experiments 5 and 6; Klein & Loftus, 1988). Given the similarities between imagining and remembering (Schacter et al., 2012), we suggest that what actually determines whether the SRE occurs when mental imagery is used is whether the imagery “illustrates” a progressive, narrative sequence of events (be it remembered or imagined), because this aspect of
episodic representation requires semantic processing\(^1\), and it seems that the SRE is largely due to highly organized and elaborate semantic processing related to the self (Klein & Loftus, 1988; Symons & Johnson, 1997). Those studies that failed to demonstrate the SRE simply had participants generate images, in some cases rather bizarre ones that would not be easily integrated into existing semantic knowledge. Indeed, Lord’s (1980, 1987) argument was that certain types of mental imagery engage “iconic” (or perceptual, graphic) encoding independent of semantic encoding, and, as mentioned earlier, recent work suggests that semantic memory is needed to provide scaffolding for episodic details when imagining events (Irish et al., 2012). In short, we think it remains reasonable to hypothesize that the PCE, which involves vividly imagining a realistic narrative, is attributable in some degree to the SRE, and to predict that removing the aspect of self-reference will diminish the PCE.

Thus, another goal of the present study is to determine whether the PCE depends on self-referential processing. We will use a modified version of Gaesser & Schacter’s (2014; Experiment 1) original paradigm, adding a third condition (imagine other) in which participants are asked to imagine another person, rather than themselves, helping people in need.

1.3 Overview of the Present Study (Experiments 1 and 2)

While previous work has shown that vividly imagining oneself helping individuals in need increases one’s self-reported willingness to help those individuals (Gaesser & Schacter, 2014), it remains unclear how much this effect depends on self-referential processing. Also, while subjective ratings of the vividness of imagined events were found to be predictive of willingness to help, such ratings do not distinguish the influence of episodic detail from that of semantic detail. The present study addresses both of these issues.

We addressed the issue of self-referential processing first (i.e., in our Experiment 1), because doing so required very little modification of the paradigm used by Gaesser & Schacter (2014).

---

\(^1\) Individuals with TLE, who have deficits in vivid episodic recollection but whose semantic memory is relatively spared, have intact narrative abilities, evidenced by their normal performance on a script generation task (St-Laurent et al., 2011). Moreover, people with semantic memory deficits have trouble imagining (Irish et al., 2012).
This allowed us to more closely replicate their original findings before adapting the paradigm to the AI method (in our Experiment 2) in pursuit of the episodic detail question.

In Experiment 1, we showed participants a series of vignettes describing individuals in crisis situations and, for each vignette, we asked them to either 1) vividly imagine themselves helping the person in need (imagine self condition); 2) vividly imagine someone else helping the person in need (imagine other condition); or 3) solve math problems unrelated to the vignettes as a baseline activity (math condition). Later, participants were shown each story again, and this time provided subjective ratings of the vividness of their imagined helping scenarios and of other phenomenological qualities. We predicted 1) that participants’ ratings of willingness to help would be highest in the imagine self condition and lowest in the math condition; 2) that these ratings would reach an intermediate level in the imagine other condition; and 3) that they would correlate with subjective ratings of vividness.

In Experiment 2, we adopted a similar design, but did not include an imagine other condition, and participants solved crossword puzzles (instead of math problems) as a control condition. Also, after imagining themselves helping, participants orally described their imagined events, which we recorded and analyzed as per the AI. We predicted 1) that participants’ ratings of willingness to help would be higher in the imagine self vs. the crossword condition; 2) that these ratings would correlate with episodic detail content in transcripts of imagined helping events; and 3) that they would correlate with subjective ratings of vividness.

2 Method: Experiment 1

2.1 Participants

Forty-four healthy, young adults (30 females, age range: 18-30 years old, $M = 21.33, SD = 3.73$) were recruited either through an online experiment database associated with an introductory psychology course at the University of Toronto or from the surrounding community by telephone or poster. Our inclusion criteria required participants to be fluent in English, to have had no prior diagnosis of a psychiatric or neurological condition, and to have normal or corrected-to-normal hearing and vision. After they completed the experiment, participants were compensated either with two course participation credits (if they were enrolled in the psychology course) or with $20
in cash. The experiment protocol was approved by the University of Toronto Research Ethics Board.

Two participants’ data have been excluded from all analyses. In one case the participant’s data failed to be recorded (due to a software error), and, therefore, could not be analyzed. In another case, data were excluded because the participant appeared to have had singular difficulty imagining vivid scenarios, yielding mean ratings of perceptual detail and spatial coherence that, in most conditions, fell more than two standard deviations below the sample mean. Of particular concern was that this person’s mean rating of spatial coherence in the imagine self condition fell more than three standard deviations below the sample mean (this was used as a conservative cut-off score for outliers, and no other participant met this cut-off). Thus, after excluding these two participants, our final sample size for all analyses was 42.

2.2 Materials and Design

2.2.1 Stimuli

Forty single-sentence vignettes served as the bases for participants’ imaginary scenarios. Each vignette featured an anonymous and nondescript individual in a specific situation of need (see Appendix A for examples). These stimuli were borrowed with permission from another lab (see Gaesser & Schacter, 2014). One of the authors (C.S.) produced a set of 96 mathematical word problems for participants to solve in the math condition. The problems described simple transactions among individuals, using names from both genders and from various ethnic backgrounds (see Appendix A for examples). Vignettes and math problems were presented one at a time on a computer screen, in black text against a grey background, using E-Prime 2.0 (Psychology Software Tools Inc. Release candidate version 2.0.10.242). Vignettes appeared in the centre of the screen for 10 seconds each. This was followed by a brief message indicating which task the participant was to perform on that trial. On imagine self and imagine other trials, the screen then became blank for 60 seconds. On math condition trials, the participant was given 60 seconds to solve as many math problems as possible, and a given math problem would remain in the centre of the screen indefinitely until either the participant pressed a key to proceed to the next problem or the 60-second time limit for the trial had elapsed. At the end of every trial, the computer played a bell sound and the screen was blank for 2.5 seconds to give the participant time to prepare for the next vignette.
2.2.2 Design

The experiment had a within-participant design, with the independent variable being the task that participants performed (3: imagine self; imagine other; math) after reading each vignette. For each participant, the vignettes were randomly assigned to each of the three task conditions and to an initial practice condition. This assignment was counterbalanced such that, across participants, each vignette was used several times in each task condition and in the practice condition. For a given participant, each of the three task conditions contained 10 different vignettes, for a total of 30 trials. These were pseudo-randomly shuffled with the restriction that no two consecutive trials belong to the same task condition. The initial practice session contained a separate set of six vignettes. A colleague had recommended we include practice trials to ensure participants thoroughly understood and were comfortable with vividly imagining hypothetical scenarios (B. Gaesser, personal communication, December 19, 2014), and we found in our own piloting that a practice session was indeed necessary. The remaining four vignettes (of the 40) were always presented after the participant had finished the main experiment, in connection with a separate pilot investigation not reported in this paper.

The primary dependent variable in our study was the participant’s willingness to help an individual in a situation of need, measured by self-report using a 7-point rating scale (1 = not at all willing; 4 = moderately willing; 7 = very willing). Participants provided other ratings as well, and these are described in the Procedure section.

2.2.3 Instruments

2.2.3.1 Toronto Empathy Questionnaire (TEQ)

In order to control for the potential influence of trait empathy on our primary dependent variable (willingness to help), we administered the 16-item TEQ, which was recently developed in order to tap into a single empathy factor shared by a wide variety of empathy measures (Spreng, McKinnon, Mar, & Levine, 2009). The TEQ has high internal consistency (Cronbach’s $\alpha = .85$) and test-retest reliability ($r = .81$, $p < .001$). It also correlates well with behavioural measures of Theory of Mind (ToM), such as the “Reading the Mind in the Eyes” Test (MIE; Baron-Cohen, Wheelwright, Hill, Raste & Plumb, 2001).
2.2.3.2 Interpersonal Reactivity Index (IRI)

We also administered the Interpersonal Reactivity Index (IRI; Davis, 1980), which is an older and more commonly used measure of trait empathy. It contains 28 items, comprising four subscales of seven items each: Perspective-Taking (PT) taps into the tendency to adopt the psychological point of view of other individuals (e.g., “I try to look at everybody's side of a disagreement before I make a decision”); Fantasy (FS) assesses a person’s imaginative ability to identify with characters in books and films (e.g., “I daydream and fantasize, with some regularity, about things that might happen to me”); Empathic Concern (EC) gauges how often one feels sympathy and concern for others in unfortunate circumstances (e.g., “When I see someone being taken advantage of, I feel kind of protective towards them”); and Personal Distress (PD) is intended to probe one’s uneasiness, anxiety, and general arousal level in tense and emotionally-charged situations (e.g., “When I see someone who badly needs help in an emergency, I go to pieces”; Davis, 1983).

2.3 Procedure

Upon entering our lab, participants gave written consent to be involved in the study and then filled out an intake form providing basic demographic information. Next they were seated in a room with normal lighting, at a desk, in front of a computer, where the experimenter gave instructions regarding all tasks to be undertaken.

2.3.1 Instruction phase

Instructions were as follows:

Welcome to the experiment! In this study, we will ask you to read a series of short stories on the computer, one at a time, and we will examine how each story affects your performance on other activities in the experiment. These stories involve individuals in situations of need, and they are based on true events that people have posted about on media websites (all personally identifying information has been removed so the persons in the stories can remain anonymous). Each story is only one sentence long, and you will have 10 seconds to read it. When the 10 seconds are up, instructions will appear on the screen asking you to spend 60 seconds performing one of these activities: 1) Solve math
problems; 2) Imagine yourself helping the person in the story; 3) Imagine someone else helping the person in the story.

2.3.1.1 “Math” instructions

In the math activity, mathematical word problems will appear on the screen one at a time. You can use the pen and answer sheets provided to record your answer to each question. For the first math question you see on the computer, write down your answer on the paper in the box labelled, “1”, then write your answer to the second question in the box labelled, “2”, and so on. Each time you answer a question, press ENTER on the keyboard and the next question will appear on the screen. Try to answer as many math questions as you can, without sacrificing accuracy for speed. It is OK to guess, but first try your best to give the correct answer. Feel free to use the blank space on the answer sheets to work out the math problems.

2.3.1.2 “Imagine yourself helping” instructions

In this activity, try to vividly imagine an event in which you are interacting positively with the person in need and are helping him or her to solve the problem. The event that you imagine should take place at a specific time and location (e.g., next Tuesday evening, at the gym). It should be at least a few minutes long, and no longer than a day. You want to imagine a practical way in which you yourself could actually help the person in real life. The person you are helping should not be anyone you know personally; he or she should be a complete stranger. Aside from that, you can imagine any sort of person you like. Try to picture the event clearly in your mind's eye and make it as elaborate and as detailed as possible. For example, try to imagine what the scene looks like, what the person in need looks like, what you might say to each other, and any thoughts or feelings you both might have at the time. It is also important to imagine perceptual details even if they are not central to the story. For example, if you imagine an event taking place in a living room, picture what is on the walls (e.g., flowery wallpaper, plain white paint), any furniture in the room, and even what the floor is like (e.g., pink shag carpet, marble tile). During this activity, the screen will be blank. You should close your eyes because it will make it easier to imagine the scene more vividly (a bell sound will play to let you know when the time is up).
2.3.1.3 “Imagine someone else helping” instructions

The instructions for this condition were exactly the same as those for the imagine self condition, except for the stipulation that participants imagine the experimenter helping the person in need and that they not include themselves in the imagined scenario at all. The reason we used the experimenter as the “other person helping” was to ensure that all participants would be imagining the same person (and one of whom they had no prior knowledge).

2.3.2 Task phase

After giving the instructions, the experimenter provided an overview of what would be involved in subsequent stages of the experiment, making it clear to participants that they would later need to answer questions on the computer about the vignettes, their reactions to them, and what they did during the 60 seconds after reading each vignette. Then the experimenter guided participants through a series of practice trials to ensure that they understood and were comfortable imagining hypothetical scenarios.

2.3.2.1 Practice trials

On a practice trial, participants would spend 60 seconds imagining themselves helping the person in need and then describe out loud, and in as much detail as possible, the scenario that they had imagined. Then the experimenter would “coach” the participants (i.e., provide feedback regarding the types of details generated in the imagined event and the narrative quality of the scenario) until they felt they understood and were comfortable with the imagination task.

2.3.2.2 Experimental (“real”) trials

When participants were ready to begin, the experimenter left the room and allowed them to work in silence as the computer guided them through 30 trials (10 imagine self, 10 imagine other, and 10 math trials intermixed). On a given trial, a single-sentence vignette would appear in the centre of the screen for 10 seconds, followed by brief instructions depending on which condition the vignette had been assigned to (imagine self / other: “Imagine an event in which you / the experimenter help this person. Try to make it as vivid as possible”; math: “Solve as many math problems as you can. Try to be as accurate as possible”). The instructions remained on the screen until participants pressed the ENTER key to indicate they were ready to begin the trial. Then, for 60 seconds, the screen either became blank (in the imagine conditions) or displayed a series of
mathematical word problems (in the *math* condition). Math problems were shown one at a time and remained on the screen until participants pressed the ENTER key to indicate they were ready for the next math problem. Participants were given a pen and answer sheets on which to record their answers to the math questions. Whenever a 60-second trial was finished, a bell sound would play on the computer. The screen would be blank for 2.5 seconds, and then the next vignette would appear. Altogether, the instruction phase, the practice phase, and the 30 experimental trials typically took about one hour to complete.

### 2.3.3 Question phase

Next, all 30 trial vignettes were displayed again, one by one, but this time participants answered a series of questions on the computer after each vignette. Participants used 7-point rating scales to indicate their present willingness to help the person in need; the degree to which they experienced various emotions (*compassionate, distressed, disturbed, sympathetic, tender, warm*); and their confidence in their ability to help the person in a practical way (1 = not at all confident; 7 = very confident). Participants also rated the vividness of each imagined event, first in terms of perceptual detail: “The imagined scene in your mind was?” (1 = vague; 7 = detailed) and then in terms of spatial coherence: “The imagined scene in your mind was?” (1 = fragmented; 7 = coherent). Before beginning the question phase, the experimenter ensured that participants understood these constructs. For explanatory purposes, perceptual detail was likened to the graphic resolution of an image of a landscape, with the high end of the scale corresponding to the level of detail one would see in a realistic painting of a landscape and the low end corresponding to that depicted in a simple cartoon of the same landscape. Our definition of spatial coherence was based primarily on terms used by Hassabis et al. (2007), with the high end of the scale indicating that one had been able to “see the whole scene in my mind’s eye” and the low end indicating that the imagined event was merely “a collection of separate images”. The experimenter also likened the imagining of a “coherent” (high end of the scale) event to “watching a movie in your head” and told participants that a “fragmented” (low end of the scale) event was, in contrast, “more like a slideshow”. The question phase was self-paced; participants had as much time as they needed to answer all questions for all vignettes.

It is important to note that participants were given the option to rate detail or coherence as “zero” if they could not remember imagining anything for the vignette in question. Theoretically,
vignettes that had been assigned to the *math* condition in particular should receive “zero” ratings for both detail and coherence, since participants were not instructed to imagine anything related to the vignette on those trials. Indeed, when Gaesser & Schacter (2014, Experiment 1) compared imagining oneself helping to the math condition, vividness ratings were not even collected for math trials. However, we decided to collect vividness ratings for all trials, because we considered it possible that participants might imagine something related to the vignette during the 10 seconds when they first read it on the screen, before receiving the instructions that let them know whether a given trial was an *imagine self*, *imagine other*, or *math* trial.

Finally, participants completed the TEQ and the IRI before being debriefed, compensated, and dismissed. The entire experiment session typically took about two hours to complete (some participants required extra time).

3 Results: Experiment 1

3.1 Willingness to help

A one-way repeated measures ANOVA was conducted to examine the effect of task type on willingness to help across *imagine self*, *imagine other*, and *math* conditions (see Figure 1). There was a significant effect of task type, $F(2, 82) = 10.19, p < .001, \eta^2_p = .199$. Post-hoc tests using the Bonferroni correction revealed that, compared to the *math* condition ($M = 4.73, SD = 1.1$), mean willingness to help was significantly higher in both the *imagine self* ($M = 5.21, SD = 0.86$) and the *imagine other* ($M = 5, SD = 0.94$) conditions ($p = .001, d = .59$ and $p = .043, d = .4$, respectively). There was a trend toward willingness to help being higher in the *imagine self* condition compared to the *imagine other* condition ($p = .052, d = .38$).

Self-reported trait empathy (TEQ score; $M = 45.73, SD = 5.58$) correlated positively with willingness to help in all three conditions (*imagine self*: $r = .561, p < .001$; *imagine other*: $r = .573, p < .001$; *math*: $r = .419, p = .006$). While the correlation between TEQ score and willingness to help in the *imagine other* condition was significantly greater than it was in the *math* condition (Pearson & Filon's $z = 1.746, p = .04$), there were no other differences between correlations ($z < 1.645$ and $p > .05$ in all cases). The *emotional concern* subscale of the IRI ($M = 19.16, SD = 4.64$) also correlated with willingness to help in all conditions, albeit not as strongly (*imagine self*: $r = .39, p = .017$; *imagine other*: $r = .421, p = .009$; *math*: $r = .332, p = .045$).
Subjective vividness ratings

A one-way repeated measures ANOVA was conducted to examine the effect of task type on each of the two subjective ratings of vividness (perceptual detail and spatial coherence) in imagined scenarios across imagine self, imagine other, and math conditions (see Table 1 for descriptives). There was a significant effect of task type on perceptual detail, $F(2, 82) = 28.74, p < .001, \eta^2_p = .412$, and on spatial coherence, $F(2, 82) = 35.06, p < .001, \eta^2_p = .461$. Post-hoc tests using the Bonferroni correction revealed that there was no significant difference in either mean perceptual detail or mean spatial coherence when comparing the imagine self and imagine other conditions (detail: $p = .426$; coherence: $p > .99$), but that both imagine conditions yielded significantly higher detail and coherence than the math condition did ($p < .001$ and $d > .8$ in all cases). While the fact that vividness ratings were not “at floor” in the math condition confirms our expectation that participants were able to construct mental scenarios in all conditions during the brief reading of the vignettes (before receiving instructions about which task to engage in), the significant differences between the imagine and math conditions indicates that the explicit instruction to engage in further imagining on imagine trials was effective.
Self-reported trait empathy (TEQ score) correlated positively with subjective perceptual detail in the *imagine self* ($r = .41, p = .008$) and *imagine other* ($r = .337, p = .031$) conditions. It also correlated positively with subjective spatial coherence in the *imagine self* condition ($r = .429, p = .005$) and there was a trend toward the same relationship in the *imagine other* condition ($r = .297, p = .06$). None of the IRI subscales correlated with subjective vividness ratings in any of the three conditions.

<table>
<thead>
<tr>
<th></th>
<th>Subjective perceptual detail</th>
<th>Subjective spatial coherence</th>
<th>Confidence in one's own ability to help</th>
<th>Feeling compassionate</th>
<th>Feeling warm</th>
</tr>
</thead>
<tbody>
<tr>
<td>imagine self</td>
<td>4.63 (.91)</td>
<td>4.42 (.92)</td>
<td>4.68 (0.96)</td>
<td>4.63 (.91)</td>
<td>3.29 (1.24)</td>
</tr>
<tr>
<td>imagine other</td>
<td>4.47 (.98)</td>
<td>4.34 (1.01)</td>
<td>4.4 (0.92)</td>
<td>4.52 (.88)</td>
<td>3.31 (1.1)</td>
</tr>
<tr>
<td>math</td>
<td>3.24 (1.38)</td>
<td>2.98 (1.42)</td>
<td>4.01 (1.04)</td>
<td>4.28 (.91)</td>
<td>2.94 (1.26)</td>
</tr>
</tbody>
</table>

*Table 1.* Means for subjective ratings associated with imagined helping scenarios, by task type, in Experiment 1. Standard deviations are given in parentheses.

### 3.3 Relationship between vividness and willingness to help

In all three conditions, subjective ratings of perceptual detail correlated with willingness to help (*imagine self*: $r = .489, p = .001$; *imagine other*: $r = .567, p < .001$; *math*: $r = .342, p = .027$), as did ratings of spatial coherence (*imagine self*: $r = .409, p = .007$; *imagine other*: $r = .501, p = .001$; *math*: $r = .373, p = .015$; see Figure 2). No two correlations were significantly different from each other (Pearson & Filon's $z < 1.645$ and $p > .05$ in all cases) although there was a trend toward perceptual detail being more strongly correlated with willingness to help in the *imagine other* condition compared to the *math* condition ($z = 1.444, p = .074$).

Further, to maximize sensitivity and statistical power, we used a multilevel modeling approach with the individual trial as the unit of analysis. This allows us to account for the fact that the vignettes vary in the degree to which they engender vivid imagination and willingness to help; it
also increases the sample size, since each trial is treated as a separate case, with the model
nesting trials within participants and accounting for the random effect of participant. Modeling
revealed that perceptual detail ratings significantly predicted willingness to help on a trial-by-
trial basis in all conditions (imagine self: $\beta = .385, N = 417, p < .001$; imagine other: $\beta = .345, N = 415, p < .001$; math: $\beta = .352, N = 415, p < .001$), as did ratings of spatial coherence (imagine
self: $\beta = .379, N = 417, p < .001$; imagine other: $\beta = .272, N = 415, p < .001$; math: $\beta = .317, N = 415, p < .001$). All models also included TEQ score, in order to account for individual
differences in trait empathy, and this was consistently found to be a significant (albeit relatively
weak) predictor alongside vividness ratings ($\beta < .1$ and $p < .025$ in all cases).

Figure 2. Mean subjective vividness ratings plotted against mean willingness to help in all three
conditions of Experiment 1, at the participant level. A) imagine self; B) imagine other; C) math.

3.4 Confidence and emotion ratings

A one-way repeated measures ANOVA was conducted to examine the effect of task type on
confidence in one’s own ability to help the person in need across imagine self, imagine other,
and math conditions (see Table 1 for descriptives). There was a significant effect of task type on
confidence, $F(2, 82) = 13.34, p < .001, \eta^2_p = .246$. Post-hoc tests using the Bonferroni correction
revealed that there was no significant difference in mean confidence when comparing the
imagine self and imagine other conditions ($p = .079$), but that both imagine conditions yielded
significantly higher confidence than the math condition did (imagine self: $p < .001$, $d = .71$; imagine other: $p = .009$, $d = .49$). Importantly (and intuitively), confidence correlated very strongly with willingness to help in all three conditions at the participant level (imagine self: $r = .786$; imagine other: $r = .775$; math: $r = .779$; all $p$’s < .001). Similarly, at the trial level, hierarchical regression revealed confidence to be a strong predictor of willingness to help in all conditions (imagine self: $\beta = .535$; imagine other: $\beta = .552$; math: $\beta = .482$; all $p$’s < .001).

Finally, one-way repeated ANOVAs were conducted on participants’ self-reported emotional reactions to vignettes (see Table 1 for descriptives). There was a significant effect of task type on feeling compassionate, $F(2, 82) = 8.84$, $p = .013$, $\eta^2_p = .101$, and on feeling warm, $F(2, 82) = 13.34$, $p < .001$, $\eta^2_p = .177$. Post-hoc tests using the Bonferroni correction revealed that participants reported feeling significantly more compassionate in the imagine self compared to the math condition ($p = .036$, $d = .41$), and significantly “warmer” in the imagine self and imagine other conditions compared to the math condition ($p = .006$, $d = .51$ and $p = .004$, $d = .54$, respectively).

4 Discussion: Experiment 1

In Experiment 1, as predicted, we found that willingness to help was significantly higher in the imagine self condition compared to the math condition, successfully replicating the basic PCE observed by Gaesser & Schacter (2014). We had also predicted that willingness to help in the imagine other condition would be higher than in the math condition, as was indeed the case. Our prediction that imagine self would have an advantage over imagine other was marginally borne out; the effect size was moderate ($d = .38$) but just shy of statistical significance ($p = .052$). Finally, subjective vividness ratings correlated with willingness to help in all conditions, as predicted, at both the participant level and the trial level.

The main question we were asking in Experiment 1 was whether the PCE is attributable, at least in some degree, to self-referential processing. We found tentative evidence that this is the case, as imagining oneself helping led to marginally greater willingness to help than imagining another person helping. Since this experiment was conducted, however, a more recent study (Gaesser, Horn, & Young, in press) has shown that imagining oneself helping does lead to significantly greater willingness to help compared to imagining someone else, and a close comparison between our results and theirs reveals virtually identical effect sizes ($\eta^2_p = .13$ and .132,
respectively). The reason that our effect was only marginally significant is probably that we tested fewer participants compared to Gaesser et al. (in press; N = 42 vs. 50).

In any case, the role of self-reference in the PCE is probably not to render imagined events more detailed and thus more memorable (as we had initially argued), given that there were no significant differences in subjective vividness ratings between the imagine self and imagine other conditions. There was also no difference in self-reported confidence in one’s own ability to help between these two conditions, which makes sense: if one is able to imagine, in detail, someone else implementing a practical solution to a problem (one that requires no expert knowledge or skill, which is arguably true of all the vignettes we used), it is likely that one actually knows how to solve the problem oneself, and is therefore just as confident in one’s own ability to solve it regardless of having imagined oneself or someone else doing so.

Importantly, though, confidence ratings were significantly higher in both imagine conditions compared to the math condition, and strongly correlated with willingness to help in all conditions. It is likely that a major mechanism by which episodic representation increases willingness to help is the bolstering of one’s confidence in one’s ability to help. We suggest that such confidence increases when appropriate evidence is made available in the form of a memory of an imagined scenario in which one helps. This is consistent with the idea presented earlier that the PCE basically operates according to the availability heuristic (Tversky & Kahneman, 1973). In other words, people decide they are more willing to help when an example of helping is readily available to their minds due to their having recently imagined it. This fits with previous work showing that imagining future events leads people to perceive them as being more plausible (Carroll, 1978; Schacter, Benoit, De Brigard, & Szpunar, 2015; Szpunar & Schacter, 2013). It is also worth noting that, when asked at debriefing on what they based their ratings of willingness to help, many of our participants said that they were more willing to help if they could conceive of themselves doing so, which is to say, if it seemed plausible to them.

But if the role of self-reference in the PCE is not to render imagined events more detailed (and thus more memorable) nor to increase one’s confidence in one’s own ability to help, then what is it? Part of the answer may be found in an examination of participants’ self-reported emotional reactions. While participants felt more compassionate on imagine self compared to math trials, there was no such difference between imagine other and math trials. Perhaps, then, while simply
imagining a specific helping strategy being carried out is enough to address the question (in the participant’s mind) of whether there is a practical means of helping the individual in need, imagining oneself in an interpersonal exchange with that individual provides an emotional motive to help, and it is because of this additional component that imagining oneself helping leads to greater willingness to help than imagining someone else helping.

On a related note, we also found that self-reported trait empathy correlated strongly with willingness to help in all conditions. TEQ scores in particular correlated with subjective vividness ratings as well (in the imagine conditions). These results gave rise to two new questions: 1) since the TEQ was not administered until the end of the experiment, were participants’ responses to it influenced by their having vividly imagined the helping scenarios?; and 2) to what degree is the TEQ itself dependent on episodic representation? To address the first question, in Experiment 2 we decided to administer the TEQ at the beginning of the experiment. The second question will be addressed later, in our presentation of Experiment 3.

Readers are reminded that in Experiment 2 our main question was whether willingness to help is related specifically to the episodic detail of imagined events (as opposed to semantic detail). This study was very similar to Experiment 1 except that participants described their imagined events orally while being recorded, and we analyzed the transcripts for their episodic detail content according to the AI procedure. We did not include an imagine other condition in Experiment 2, since it was not necessary for addressing the episodic detail question. As a control condition, participants solved crossword puzzles (after reading the vignettes) and recalled an event from their own lives related to each word they came up with for the crosswords. We chose this control condition because we wanted to control for the effects of episodic representation per se and of spending a few minutes verbally describing one’s past episodic activity.

Our main predictions for Experiment 2 were 1) that our participants’ ratings of willingness to help would be higher on imagine self vs. crossword trials; 2) that these ratings would correlate with episodic detail in transcripts of imagined helping scenarios; and 3) that they would also correlate with subjective ratings of vividness.

Our prediction for the TEQ was that, if Experiment 1 participants rated themselves as more empathetic than they normally would have because of having recently imagined helping, then
TEQ scores in Experiment 2 should not correlate with willingness to help and should be lower than in Experiment 1, since participants now write the TEQ before they imagine helping.

5 Method: Experiment 2

5.1 Participants

Twenty-five healthy, young adults (18 females, age range: 18-34 years old, $M = 20.7, SD = 3.61$) were recruited according to the same protocol as in Experiment 1. Three participants’ data have been excluded from all analyses due to either misunderstanding of, or failure to comply with, instructions (e.g., imagining helping persons in need in a control condition where participants were instructed to perform a different activity). After excluding these participants, our final sample size for all analyses was 22.

5.2 Materials and Design

5.2.1 Stimuli

As Experiment 2 required fewer trials than Experiment 1 (due to time constraints associated with the AI procedure) we used a subset of the vignettes, which are described more thoroughly in the Experiment 1 Method section. We reviewed piloting data that we had previously collected on the complete set of 40 vignettes and selected 14 that were associated with high levels of detailed imagining. Vignettes were presented one at a time on a computer screen, in black text against a grey background, using E-Prime 2.0 (Psychology Software Tools Inc. Release candidate version 2.0.10.242), as in Experiment 1.

5.2.2 Design

We adopted a within-participant design, with the independent variable being the task that participants performed ($2$: imagine self; crossword puzzle) after reading each vignette. For each participant, four vignettes were assigned randomly to each condition; this was counterbalanced across participants such that, over the course of the experiment, each vignette was used many times in each condition. This amounted to a total of eight trials per participant. There was also an initial practice session, comprised of the remaining six vignettes, as there was in Experiment 1.
The *crossword* task was intended to control for the effect of episodic representation *per se*. Thus, after reading a vignette assigned to the *crossword* condition, participants spent 60 seconds solving simple crossword puzzles (unrelated to the vignette) while conjuring up a vivid episodic memory from their own lives that they spontaneously associated with each word they came up with in answer to the crossword puzzle clues. All correct answers to clues were concrete food-related words (e.g., “apple”, “sugar”). Participants were told that if they could not think of a memory related to a given word, they should vividly imagine the item represented by the word. In both cases (remembering or imagining), the mental imagery was not supposed to be related in any way to the “helping” vignette they had read before. After 60 seconds had elapsed, participants described, out loud, what they remembered or imagined and any other thoughts they had had while working on the crossword. This was meant to control for the effect of spending a few minutes verbally describing one’s past episodic activity, which participants also did on *imagine self* trials regarding their imagined helping scenarios. Participants’ oral descriptions, on both *crossword* and *imagine self* trials, were ultimately transcribed by the experimenter and analyzed for their episodic detail content using an adapted version of the Autobiographical Interview (Levine et al., 2002). Albeit a rough comparison, because the two tasks were quite different, this allowed us to compare the *imagine self* to the *crossword* condition for the vividness of task-related mental imagery.

As in Experiment 1, the primary dependent variable was willingness to help.

### 5.2.3 Toronto Empathy Questionnaire (TEQ)

As in Experiment 1, we used the TEQ to account for individual differences in trait empathy. A description of this measure can be found in the Experiment 1 Method section. We decided not to use the IRI in Experiment 2 for two reasons: 1) the IRI took longer to administer, and we needed to keep our experiment duration at two hours or less for administrative purposes; 2) the TEQ was more strongly correlated with willingness to help in Experiment 1 than the IRI was, and the IRI did not correlate with subjective vividness ratings at all in Experiment 1 while the TEQ did. As mentioned above, we had decided to investigate the TEQ’s relationship to imagining helping (Experiment 3) and thus collected the requisite data while administering the TEQ to participants at the beginning of Experiment 2.
5.3 Procedure

The procedure for Experiment 2 was nearly identical to that of Experiment 1, except for the addition of the oral description task. Also, there was no imagine other condition and the math condition was replaced by the crossword condition.

5.3.1 Instruction phase

5.3.1.1 “Crossword puzzle” instructions

You will be provided with a set of crossword puzzles at the beginning of the experiment. Whenever you see instructions to work on the crosswords, try to complete as many as you can in the time you have. If you have not finished a crossword puzzle by the time the 60 seconds are up, you can put it aside and continue working on it the next time you see instructions to work on crosswords. Whenever you think you know what one of the words in a crossword is, take a moment to vividly recall a personal memory related to the word; close your eyes and try to picture it in your mind as clearly as you can. (If you cannot think of a memory related to the word, simply visualize the object as vividly as possible.) For example, if you think one of the words is, "museum", you should try to remember a time when you visited a museum. If you don't have any memories of museums, you would just imagine what a museum looks like. When the 60 seconds are over, a bell sound will play to let you know the time is up. Then you will be asked to describe the things you remember or imagined and any other thoughts you had while working on the crosswords (e.g., any special strategies you may have used). You will have as much time as you need to describe these things, and the experimenter will be digitally recording your response so we can analyze it later.

5.3.1.2 “Imagine yourself helping” instructions

These instructions were identical to those given in Experiment 1, except that participants were also told that they would need to describe the events they imagined in as much detail as possible, that they would have as much time as they needed to describe each event, and that the experimenter would be digitally recording their responses for later analysis.
5.3.2 Task phase

After giving the instructions, the experimenter provided an overview of what would be involved in subsequent stages of the experiment, making it clear that participants would later need to answer questions about the vignettes, their reactions to them, and any helping-related mental imagery they had experienced. Then the experimenter guided participants through a set of practice trials as in Experiment 1.

When participants were finished with the practice trials, the experimenter remained in the room and allowed participants to work in silence as the computer guided them through each of eight trials (four imagine self and four crossword trials intermixed), becoming involved only when it was time to record participants’ oral responses. On a given trial, a single-sentence vignette would appear in the centre of the screen for 10 seconds, followed by brief instructions depending on which of the two task conditions the vignette had been assigned to (imagine self: “Imagine an event in which you help this person. Try to make it as vivid as possible”; crossword: “Work on the crossword puzzles”). These instructions remained on the screen until participants pressed the ENTER key to indicate they were ready to begin the trial. Then, for a period of 60 seconds, the screen became blank. On imagine self trials, participants closed their eyes and imagined helping the person mentioned in the vignette, while on crossword trials they worked on printed crossword puzzles, pausing to recall or imagine something after each word they wrote down in the puzzle. Whenever a 60-second trial was finished, a bell sound played on the computer and the screen displayed a brief message indicating it was time for participants to provide their verbal description. The experimenter instructed participants to describe what they had imagined or recalled during the previous 60 seconds and began recording the description using a handheld digital recording device. Once participants indicated they were finished speaking, the experimenter pressed a key on the computer to advance to the next trial. The screen became blank for 2.5 seconds, and then the next vignette appeared on the screen. Altogether, the instruction phase, the practice trials, and the eight experimental trials typically took about an hour and a half to complete.

5.3.3 Question phase

The question phase was identical to that in Experiment 1. Note that, in theory, crossword trials should receive “zero” ratings for subjective detail and coherence; although participants were
asked to generate mental imagery in response to crossword puzzle words, the experimenter made it clear that the vividness ratings (indeed, all the ratings) pertained to the “helping” vignettes and not to the crossword puzzles. Participants were told that if they could not remember imagining anything at all for a particular vignette, they should give a zero rating for detail and for coherence. We collected vividness ratings for *crossword* trials for the same reason we did for *math* trials in Experiment 1, namely because we anticipated that participants might spontaneously imagine something related to the vignette even when instructed not to.

Finally, the participant was debriefed, compensated, and dismissed. The entire session typically took about two hours to complete (some participants required extra time).

### 5.4 Scoring Transcripts

Once data collection was finished, the transcripts of participants’ oral descriptions of imagined events in the *imagine self* condition were scored by the experimenter for their episodic and semantic detail content according to the guidelines laid out by Levine et al. (2002) and adapted by Sheldon et al. (2011) and others (e.g., Addis et al., 2008). The verbal material was segmented into chunks, with each chunk defined as a single observation, fact, thought, emotion, or judgment. These chunks were then classified as either *internal details* (i.e., details specific to the context of the event being imagined, such as perceptual details; references to a particular time or place; moment-by-moment happenings in the unfolding of the narrative; any thoughts or emotions imagined to be experienced by the participant or characters in the story) or *external details* (i.e., details not specific to the event being imagined, such as bits of general knowledge or factual information; judgments; repeated details). When scoring was finished, we counted the total number of internal details in each transcript and also tallied the subcategories of internal details (*time, place, perceptual, emotion-thought, event*) separately. We tallied external details but did not distinguish among subcategories of external details. In order to measure episodicity while accounting for variance in total verbal output, we computed internal detail proportions for each transcript by dividing the number of internal details (or, in most cases, the number of a subcategory of internal details) by the number of total details, including external details. For example, the proportion of *time* details was calculated using the following formula:

\[
\frac{\text{Number of time details}}{\text{(Number of all internal details + number of all external details)}}
\]
Transcripts based on mental imagery conjured up in the *crossword* condition were also scored. In cases where participants did not remember specific episodes in response to crossword clues but only visualized (or simply mentioned thinking about) static objects, details associated with such imagery were scored as external. When scoring was finished, proportions of internal details were computed for *crossword* transcripts as they were for *imagine self* transcripts.

6 Results: Experiment 2

6.1 Willingness to help

A repeated measures t-test was conducted to examine the effect of task type on willingness to help between *imagine self* and *crossword* conditions (see Figure 3). There was a significant effect of task type, \( t(21) = 4.27, p < .001, d = .91 \), with mean willingness to help being greater in the *imagine self* condition (\( M = 5.24, SD = 1.31 \)) compared to the *crossword* condition (\( M = 4.53, SD = 1.35 \)). Self-reported trait empathy (TEQ score; \( M = 44.41, SD = 6.65 \)) was not significantly correlated with willingness to help in either condition (*imagine self*: \( r = .15, p = .505 \); *crossword*: \( r = .244, p = .273 \)), which was in contrast to our Experiment 1 results. Visual inspection of scatterplots, however, revealed that a single outlier was likely responsible for the lack of correlation (see Figure 4). After removing this outlier, the correlations’ \( p \)-values remained non-significant, but the coefficients did increase dramatically in magnitude (*imagine self*: \( r = .318, p = .16 \); *crossword*: \( r = .391, p = .079 \)). An independent samples t-test compared mean TEQ score between Experiment 1 (\( M = 45.73, SD = 5.58 \)) and Experiment 2 (\( M = 44.41, SD = 6.65 \)), and no significant difference was found, \( t(61) = .838, p = .405 \).
Figure 3. Mean willingness to help by task type in Experiment 2. Error bars represent the standard error of the mean. The asterisk indicates that the difference between conditions was significant at $p < .001$.

Figure 4. Mean willingness to help plotted against TEQ score, by task type, in Experiment 2. The same outlier participant is identified with a red circle in each scatterplot.
6.2 Subjective vividness ratings

Repeated measures t-tests were conducted to examine the effect of task type on each of the two subjective ratings of vividness (perceptual detail and spatial coherence) in imagined events related to the vignettes in imagine self and crossword conditions. There was a significant effect of task type on perceptual detail, \( t(21) = 6.28, p < .001, d = 1.34 \), and on spatial coherence, \( t(21) = 6.85, p < .001, d = 1.46 \), with both measures being greater in the imagine self condition (detail: \( M = 5.69, SD = 1.09 \); coherence: \( M = 5.41, SD = 1.4 \)) compared to the crossword condition (detail: \( M = 3.7, SD = 1.59 \); coherence: \( M = 3.18, SD = 1.58 \)). While the fact that vividness ratings were not “at floor” in the crossword condition confirms our expectation that participants were able to construct mental scenarios in both conditions during the brief reading of the vignettes (before receiving instructions about which task to engage in), the significant differences between conditions indicates that the explicit instruction to engage in further imagining on imagine self trials was effective.

Self-reported trait empathy (TEQ score) did not correlate with subjective vividness ratings in either of the two conditions.

6.3 Transcript scores (episodicity)

Repeated measures t-tests were conducted to examine the effect of task type on episodicity, which we defined as the proportion of total scored details that were episodic (or “internal”), in transcripts of participants’ verbal descriptions of what they thought about during imagine self and crossword trials. Readers are reminded that such details were related to imagined helping scenarios in the imagine self condition and to other, irrelevant mental imagery in the crossword condition. There was a significant effect of task type on total internal detail content, \( t(21) = 7.2, p < .001, d = 1.53 \), such that this measure was greater in the imagine self condition compared to the crossword condition (see Figure 5). Regarding subcategories of internal detail, there was a significant effect of task type on place detail, \( t(21) = 5.03, p < .001, d = 1.07 \), perceptual detail, \( t(21) = 5.18, p < .001, d = 1.11 \), and event detail, \( t(21) = 6.48, p < .001, d = 1.38 \), with all of these measures being greater in the imagine self condition compared to the crossword condition. There was no difference between conditions when it came to time detail, \( t(21) = -1.14, p = .268 \), and emotion-thought detail, \( t(21) = -.4, p = .695 \).
We asked an independent rater (an undergraduate student in our lab, who was naïve to the purpose of the study) to score a subset of participants’ transcripts so that we could obtain measures of interrater reliability. Fortunately, agreement between the two raters was highly consistent for *imagine self* transcripts, with ICC’s ranging from .804 (for internal perceptual details) to .964 (for internal event details). For *crossword* transcripts, however, agreement was relatively inconsistent, with ICC’s ranging from -.209 (for internal place details) to .887 (for internal event details), and this is probably because the content in *crossword* transcripts was quite heterogeneous and in many cases unclear (e.g., while some participants recalled specific personal memories related to words from the crosswords, others simply visualized the items, or related bits of general knowledge of the items, or did some combination of the above, in a sort of “stream-of-consciousness” fashion).

![Figure 5](image-url)  
*Figure 5.* Mean internal details as proportions of total details (total number of internal + total number of external) by task type in Experiment 2. Error bars represent the standard error of the mean. An asterisk indicates that the difference between conditions was significant at $p < .001$. 
6.4 Relationships among vividness ratings, internal detail, and willingness to help

6.4.1 Subjective vividness ratings

Subjective ratings of perceptual detail and spatial coherence were highly correlated with each other in both conditions (imagine self: $r = .91, p < .001$; crossword: $r = .908, p < .001$). In the imagine self condition, neither of these measures correlated with willingness to help ($p > .6$ in both cases). Both of them, however, correlated positively with willingness to help in the crossword condition (detail: $r = .474, p = .026$; coherence: $r = .442, p = .04$; see Figure 6 for scatterplots). In order to maximize sensitivity and statistical power, as we did in Experiment 1, we used a multilevel modeling approach with the individual trial as the unit of analysis. Modeling revealed that subjective ratings of perceptual detail significantly predicted willingness to help on a trial-by-trial basis in the crossword condition ($\beta = .241, N = 87, p = .015$), but not in the imagine self condition ($p = .592$). Parallel results emerged for subjective ratings of spatial coherence (crossword: $\beta = .241, N = 87, p = .026$; imagine self: $p = .937$). We did not include TEQ score in any of these models because, as mentioned above, it did not correlate with willingness to help at the participant level to begin with.

Figure 6. Mean subjective vividness ratings plotted against mean willingness to help in both conditions of Experiment 2, at the participant level. A) imagine self; B) crossword.
6.4.2 Internal detail content in transcripts

Given that the transcript scoring was intended to be another measure of vividness (only a more objective one), we should expect internal detail scores to correlate with subjective ratings of vividness. It is important to note, however, that we can only really expect such a correspondence to occur in the imagine self condition, where both subjective vividness ratings and transcripts are concerned with participants’ imagined experiences of helping persons in need. In the crossword condition, subjective vividness ratings also pertain to any mental imagery involving the persons in need (although we are assuming that, in this condition, they were spontaneously imagined), but the transcripts describe things that the participants recalled or imagined inspired by the words they thought of as answers to crossword clues and not related to the “helping” vignettes. Thus, in the crossword condition, we cannot expect internal detail content in transcripts to correspond to vividness ratings.

Surprisingly, we found no significant correlations in the imagine self condition between either subjective perceptual detail or subjective spatial coherence and any category of internal detail, although there was a trend toward subjective spatial coherence being positively correlated with a combined measure of internal place and perceptual details ($r = .415, p = .055$; see Table 2 for a complete correlation matrix).

At the participant level, willingness to help in the imagine self condition did not correlate with any category of internal detail ($p > .3$ in all cases). It is possible, however, that our participant-level analyses were not sensitive enough or did not have sufficient statistical power. Hence, we used a hierarchical regression approach with the individual trial as the unit of analysis, nesting trials within participants (the same sort of modeling used above to predict willingness to help with subjective vividness ratings). Internal place details significantly predicted greater willingness to help in the imagine self condition ($\beta = 7.513, N = 87, p = .029$), while internal time details significantly predicted less willingness to help ($\beta = -13.774, N = 87, p = .043$). As with the subjective vividness models documented above, we did not include TEQ score in any of the internal detail models because it did not correlate with willingness to help at the participant level. Also, we did not examine the relationship between internal detail content and willingness to help in the crossword condition; since transcripts generated on crossword trials were based on
mental imagery completely unrelated to the “helping” vignettes, we had no reason to expect such a relationship.

<table>
<thead>
<tr>
<th>Internal detail category</th>
<th>Subjective perceptual detail</th>
<th>Subjective spatial coherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>-.167</td>
<td>-.182</td>
</tr>
<tr>
<td>Place</td>
<td>.073</td>
<td>.139</td>
</tr>
<tr>
<td>Perceptual</td>
<td>.313</td>
<td>.359</td>
</tr>
<tr>
<td>Emotion-thought</td>
<td>-.253</td>
<td>-.245</td>
</tr>
<tr>
<td>Event</td>
<td>.026</td>
<td>-.034</td>
</tr>
<tr>
<td>Combined place and perceptual</td>
<td>.349</td>
<td>.415</td>
</tr>
<tr>
<td>Total internal details</td>
<td>.338</td>
<td>.347</td>
</tr>
</tbody>
</table>

*Table 2.* Correlation matrix comparing subjective ratings of vividness to internal detail scores in transcripts of imagined helping scenarios in Experiment 2.

7 Discussion: Experiment 2

In Experiment 2, as predicted, participants were more willing to help in the *imagine self* vs. the *crossword* condition. While subjective vividness ratings showed the same difference between conditions, they did not correlate with willingness to help on *imagine self* trials (contrary to our prediction) but only on *crossword* trials. Vividness ratings did not correlate significantly with internal details, although the correlation between subjective spatial coherence and a combined measure of place and perceptual detail approached significance. We found partial support for our prediction that episodic (internal) detail in transcripts from *imagine self* trials would be related to willingness to help, as internal place detail was a strong predictor in this regard. Internal time detail, however, predicted lower willingness to help, which we did not expect. Finally, when we administered the TEQ at the beginning of the experiment (instead of at the end), this measure of
self-reported trait empathy did not correlate significantly with willingness to help, nor with subjective vividness ratings.

First, the advantage of the *imagine self* over the *crossword* condition indicates that willingness to help is not increased by episodic representation *per se*. While auxiliary to our main question in Experiment 2, regarding the role of episodic vs. semantic detail, this finding is important in light of other work that implicates a common brain network involved not only in remembering and imagining, but in other “self-projection” tasks such as moral decision-making and Theory of Mind (Buckner, Andrews-Hanna, & Schacter, 2008). It is conceivable that merely activating or “priming” this brain network (e.g., by asking participants to recall memories associated with words in the crossword puzzles) could increase willingness to help. Our finding indicates, however, that this is not the case.

Second, and more generally, our basic replication of the PCE (i.e., greater willingness to help on *imagine self* vs. *crossword* trials) speaks to the robust nature of the effect. It was demonstrated several times by Gaesser & Schacter (2014), and we have replicated it twice in the present study (first in Experiment 1 and again in Experiment 2).

### 7.1 On subjective vividness ratings

In Experiment 2, however, we did not replicate the PCE in its entirety—our results clashing with those from Experiment 1—as subjective vividness ratings did not correlate with willingness to help on *imagine self* trials but only on *crossword* trials. Our first intuition was that this was due to a lack of variance in vividness ratings in the *imagine self* condition. This hunch was corroborated by visual inspection of scatterplots (see Figure 6) and formally confirmed by a Levene’s test of homogeneity of variance, which showed vividness ratings were significantly more variable in the *crossword* compared to the *imagine self* condition (subjective perceptual detail: $F = 27.302, p < .001$; subjective spatial coherence: $F = 5.624, p = .019$).

The question we must address, then, is what caused this heterogeneity, and the answer may be that participants gave their vividness ratings after they had already described their imagined events aloud and in great detail. It has long been known that elaborating on verbal material improves one’s memory for it (e.g., Craik, 2002; Craik & Tulving, 1975), and Drivdahl & Zaragoza (2001) have shown that elaborating on perceptual information in “suggested”
(hypothetical) events leads people to falsely recall those events as having actually occurred. In our Experiment 1, where vividness ratings did correlate with willingness to help in the imagine conditions, participants made their ratings long after they initially imagined the events. By that time, some of their memories of their imagined events would have deteriorated somewhat due to both decay and interference (e.g., Altmann & Gray, 2002), with a large proportion of said interference attributable to the greater number of trials in Experiment 1 vs. Experiment 2 (30 in total vs. eight in total). Presumably, then, if participants were able to remember some imagined events more easily than others, they would rate those events as more vivid than the others, giving rise to variability in the ratings. In Experiment 2, however, participants elaborated on their imagined events immediately after they produced them; they also imagined fewer events and therefore did not need to remember as many when the time came to make their ratings. These two factors made the imagined events easier for participants to remember—and, by extension, rate the vividness of—during the question phase, resulting in less variability in the ratings. Note that subjective vividness ratings on imagine self trials were also higher, on average, in Experiment 2 than in Experiment 1 (detail: $M = 5.69$ vs. 4.63; coherence: $M = 5.41$ vs. 4.42), which is what we would expect if participants had better memory for their imagined events.

The reason these things did not occur in the crossword condition is that the transcripts therein were based on mental imagery associated with the crossword puzzles, while only the vividness ratings were based on the vignettes. Thus, participants’ oral descriptions on crossword trials only strengthened their memory for their “crossword imagery” and not for any helping-related imagery they may have spontaneously generated.

To prevent this problem in future work, we could have participants make their vividness ratings before describing their imagined events aloud (e.g., St-Laurent, Moscovitch, Jadd, & McAndrews, 2014), which would avoid elaboration effects, and we could have a greater number of trials, which would allow for greater influence of interference and decay on participants’ memories of imagined events. Then, the variability of the strength of participants’ memories for their imagined events would be amplified (i.e., more vividly imagined events would likely be much easier to recall), resulting in greater variability in vividness ratings.

On a related note, there are other unexpected findings of ours that warrant explanation, although they are not contrary to any of our predictions. In both Experiments 1 and 2, subjective vividness
ratings in control conditions (i.e., *math* and *crossword*, respectively) correlated with willingness to help. By contrast, Gaesser & Schacter (2014) did not find such a correlation in their “web page” control condition: when participants vividly imagined a web page with text describing strategies for helping people (without imagining the strategies themselves being carried out), their ratings of the vividness of the web page did not correlate with subsequently reported willingness to help. Why, then, did we find such correlations in our control conditions? Probably because, in Gaesser & Schacter’s (2014) web page condition, participants were told explicitly to imagine something other than an actual helping scenario, and that was what they rated the vividness of. In our control conditions, vividness ratings were based on helping scenarios that we anticipated participants might spontaneously imagine without being asked to. Our *math* condition was identical to the one in Gaesser & Schacter’s (2014) first experiment; although their participants may have spontaneously imagined helping as well, they had no vividness ratings in their *math* condition. We collected vividness ratings in all conditions in both experiments, and consistently found *imagine* conditions to yield significantly greater vividness ratings, on average, than *math* or *crossword* conditions. What are we to conclude from all of this, then? First, since vividness ratings were not “at floor” in our control conditions, we conclude that—as we anticipated—participants spontaneously imagined helping in those conditions, probably during the initial reading of the vignette, before they received instructions about which task they were to perform. Second, since vividness ratings were significantly higher on *imagine* vs. *math* and *crossword* trials, we conclude that our explicit instructions to imagine helping on those trials were effective in prompting further imagining. Third, since vividness ratings correlated with willingness to help even in control conditions, we conclude that to the extent a person imagines helping at all, the vividness of that imagery will increase his or her willingness to help, which is in line with our predictions and with those of Gaesser & Schacter (2014).

### 7.2 On transcripts and internal details

We had also predicted that, in Experiment 2, proportions of episodic (internal) details in transcripts of imagined helping events would correlate with willingness to help. We found that only the proportion of internal place details predicts greater willingness to help, and it is puzzling why this is not true of any other type of internal detail. Even more puzzling is the finding that internal *time* detail predicts less willingness to help.
First, it should be noted that internal detail content was not significantly related to subjective vividness ratings (although a combined measure of place and perceptual detail correlated at trend level with subjective spatial coherence). This is surprising, as experience and unpublished data from our own lab and from that of Gilboa et al. (2004) have established that subjective vividness ratings generally correspond to internal detail scores in transcripts of remembered events, and there is ample evidence that both remembering and imagining events rely on many of the same cognitive and neural processes (e.g., Schacter et al., 2012). It is conceivable that, in the case of memory, verbally reported details more accurately convey the bits of information that a person’s hippocampus is integrating into his or her mental representation of a given event, compared to imagination. More likely, though, is that our participants felt pressure to provide an overabundance of detail associated with their imagined events to the point where they said much more than was required of them, so that any truly imagined information in their transcripts was “diluted” by extraneous elaboration. In our instructions, we strongly emphasized the importance of giving as much detail as possible, and this may have led to such “demand characteristics” (Orne, 2009).

Nevertheless, we found that the proportion of internal place details in transcripts of imagined helping scenarios significantly predicted willingness to help. First, what do place details represent? A place detail is assigned whenever information in a transcript indicates that a person’s imagined event is “set” in a particular spatial context (e.g., “my neighbour’s house”), but also whenever the person explicitly describes moving around within that spatial context (e.g., “I heard a noise, so I walked into the kitchen”; “I ran upstairs, went down the hall, and entered the bathroom to grab a box of bandages from the medicine cabinet”). What is the connection between such concrete references to spatial context and willingness to help? It could be that, in the “crisis scenarios” we used in our study, effectively helping the people in need requires a lot of movement in space, in which case the more vividly one imagined helping, the more prominent spatial details would be. As an informal investigation of this idea, two of the authors (C.S. and M.P.M.) reviewed the vignettes and coded them for how much they thought effective problem-solving in each situation depended on traveling between locations or generally moving about in space. Of the eight vignettes used in Experiment 2, five were judged by both authors to meet this criterion (e.g., “This person is quite sick and is having difficulty getting around their house”, “This person is suffering from dementia and is lost in a mall”). Thus, we tentatively suggest our
internal place detail finding is evidence that imaging oneself helping others in a more vivid (and “physically involved”) way renders one more willing to help those people, as we predicted.

But if this were true, would we not expect other types of internal detail to predict willingness to help as well? And how does all this relate to the aforementioned “diluting” influence of demand characteristics? A very interesting possibility is that place details were less affected by demand characteristics than other details were, because of their comparatively inconspicuous—yet vital—role in episodic representation: establishing the spatial context in which events are played out. According to scene construction theory (SCT), and based on a review of evidence from neuropsychological and neuroimaging studies, the primary function of the human hippocampus in memory and imagination is to generate scenes as an initial backdrop and foundation, “allowing the event details of episodic memories and imagined future experiences to be martialed, bound and played out in a coherent spatial context” (Mullally & Maguire, 2013, p. 1181). Consider, then, that participants in our study began their descriptions of imagined events by establishing the spatial context, and then “played out” and elaborated on events that transpired within that space, and that it was in this second step in particular that they provided and overabundance of details (due to demand characteristics). This idea is supported by the greater numbers of event and perceptual details (n = 1358 and 803, respectively), compared to place details (n = 218), across all imagine self transcripts. Granted, place details were assigned not only when participants established the initial spatial context, but also whenever they imagined moving to a new spatial context. This is appropriate, however, given that imagined events (and memories) commonly involve changes in environment (e.g., moving to a different room in a house; starting out in a hardware store but then walking to a nearby grocery store). Any elaboration, then, of things that happened within that new spatial context, or of what it looked like, would come after establishing the change in spatial context and would be scored in terms of event or perceptual details. In sum, our argument is that although 1) demand characteristics induced our participants to provide an overabundance of verbal material, diluting the details that directly corresponded to what they actually imagined, 2) place details were not diluted in this way, because they represent the “first step” of establishing spatial context, and any further description of the spatial context or the happenings therein were scored as either perceptual or event details. Hence, in the present study, place detail content is the most reliable index of what participants actually imagined, and our finding that it significantly predicted willingness to help.
in the *imagine self* condition indicates that the more vividly one imagines helping others, the more willing one will be to help them.

The inverse correlation between time detail and willingness to help is probably less meaningful. Time details were by far the least frequent type of detail. The total number of time details, across all *imagine self* transcripts from all participants, was 57. By comparison, the total number of place details was 218, and of perceptual details, 803. Across all *imagine self* transcripts, over a third of our participants (N = 8 of 22) did not mention a single time detail (full sample $M = 2.59$, $SD = 3.11$), while every participant mentioned multiple place details ($M = 9.91$, $SD = 4.34$) and perceptual details ($M = 36.5$, $SD = 19.62$). In all likelihood, there are simply insufficient data on time detail, and the inverse correlation arose due to chance or participant-related idiosyncrasies. Consistent with this suspicion, the regression coefficient was unusually large ($\beta = -13.77$), and there is no conceivable reason why time detail would exert such an extreme negative influence on willingness to help.

### 7.3 On self-reported trait empathy (TEQ scores)

In Experiment 2, we administered the TEQ at the beginning of the experiment instead of at the end, as in Experiment 1. This decision was made after finding strong positive correlations between TEQ scores and willingness to help in Experiment 1. Our intuition was that TEQ scores had been inflated in Experiment 1, because participants had been imagining themselves helping people (and thus, perhaps, perceiving themselves as more empathetic by nature) before filling out the survey, and that this also resulted in the strong correlations.

Consistent with this intuition, in Experiment 2 we did not find significant correlations between TEQ scores and willingness to help in either the *imagine self* or the *crossword* condition. Removal of a single outlier, however (who reported high willingness to help despite low trait empathy), dramatically increased the magnitude of the correlation coefficients, and visual inspection of the scatterplots (see Figure 4) revealed a trending positive relationship between willingness to help and TEQ score. Also, there was no significant difference in mean TEQ score between experiments, although it was slightly higher in Experiment 1.

Therefore, there is little or no evidence that self-reported trait empathy was influenced by imagining helping. The discrepancy in correlational results is probably due to a lack of statistical
power; our sample size in Experiment 2 was nearly half that of Experiment 1 (N = 22 vs. 42, respectively). We can conclude, simply, that people with higher trait empathy tend to be more willing, at least in hypothetical crisis situations, to help individuals whom they do not know personally (though this effect is a rather modest one).

We had also raised the question of whether the TEQ depended on episodic representation after finding strong correlations between TEQ scores and subjective vividness ratings in the imagine conditions in Experiment 1. In Experiment 2, TEQ scores did not even come close to correlating with vividness ratings, but this may be due to the lack of variance in vividness ratings in the imagine self condition, which we discussed above. Thus, it remains a possibility, based on Experiment 1 results, that people’s TEQ responses are related to the vividness of their episodic representations. Our next experiment explores this relationship in more depth.

8 Introduction: Experiment 3

Previous work has suggested that the reason imagining oneself helping increases one’s willingness to help is that it activates the self-concept, and assuming that a person’s self-concept maintains that he or she is a “moral” person, the person will respond accordingly (Gaesser et al., in press; see also Conway & Peetz, 2012; Young, Chakroff, & Tom, 2012). Consider that when a person fills out a self-report measure of trait empathy, such as the TEQ, something similar occurs. Answering questions about one’s own empathetic behaviour will activate one’s self-concept, likely bringing to mind memories of past occasions on which one has acted empathetically. Memories that are more vivid, then, may have a greater influence on one’s responses to the TEQ. Also, the relationship between willingness to help and vividness in imagined events may be partly due to a more general link between empathy and good memory for details (e.g., Ciaramelli et al., 2013). This is consistent with our finding that vividness ratings correlated with willingness to help even in baseline conditions, in which people were not asked to imagine helping but may have spontaneously done so upon first reading the vignettes.

For these reasons, we developed a “memory follow-up survey” (MemFUS) for the TEQ, predicting that if participants were to fill out the TEQ and were later asked which TEQ items brought to mind specific memories of themselves acting empathetically, we would find that participants had given themselves higher empathy ratings on those very items.
9 Method: Experiment 3

9.1 Participants

The participants in Experiment 3 were the same twenty-five healthy, young adults recruited for Experiment 1; demographic data and other information can be found above in the Participants section for that experiment. We excluded two participants simply because they did not report a single memory associated with any of the TEQ items. Thus, our final sample size for all Experiment 3 analyses was 23.

9.2 Materials and Design

9.2.1 MemFUS

One of the authors (C.S.) wrote a “memory follow-up survey” (MemFUS) for the TEQ. The MemFUS was 16 pages long (one page for each item on the TEQ; see Appendix B for an example). At the top of each page was the text of one of the TEQ items, and this was followed by a brief survey. Participants first had to answer the following question:

Did you think about a memory from your own life when you were making a rating for the above statement on the survey you were filling out earlier? If you did not, that is okay. You do not have to think of one now; you can simply skip to the next page of the survey.

If participants checked the box marked, “yes”, then they would go on to indicate whether, in the memory that the TEQ item had brought to mind earlier, they, “acted in a way that fits with the statement”, or “acted in a way that does not fit with the statement”. Finally, participants made subjective vividness ratings (in the same way they did for Experiments 1 and 2) based on the memory that had come to mind earlier (i.e., when they were filling out the TEQ), and were invited to write down a brief description of the memory (participants were told that this last part was entirely optional, since it was concerned with events from their personal lives).

9.2.2 Design

This study had a repeated measures design. The quasi-independent variable was whether or not a given TEQ item had cued, for each participant, an episodic memory of an event in which he or she was acting in a way that was congruent with the behaviour described by the TEQ item. The
dependent variable was the rating (on a scale from 0 to 4) that participants gave the item. For each participant, we computed the average TEQ score for all “cued-and-congruent” (C&C) items and a separate average for all other items.

9.3 Procedure

Participants came into our lab primarily for Experiment 1. After filling out the TEQ, participants were given the MemFUS without having been informed of it beforehand. The experimenter first collected the completed TEQ and then explained how to fill out the MemFUS. Once it was clear that participants understood what to do, the experimenter left the room and participants filled out the MemFUS (without having access to their completed TEQ). Then the experimenter reviewed participants’ responses with them and clarified any ambiguities. For example, in some cases, the memories that participants reported being cued by TEQ items were not drawn from specific, individual episodes but rather from general knowledge and experience. The latter two categories are considered to reflect semantic rather than episodic memory processes, and in the present experiment we were interested in the unique contribution of episodic memory retrieval to TEQ performance.

10 Results: Experiment 3

First, participants had nearly twice as many TEQ items that were not marked C&C ($M = 10.3$, $SD = 2.6$) as those that were marked C&C ($M = 5.7$, $SD = 2.6$), and this difference was significant, $t(22) = 4.25$, $p < .001$, $d = .89$. Thus, we used a non-parametric test to assess the difference in TEQ score between these two categories of items. A related samples Wilcoxon signed rank test revealed that the median TEQ score for items marked C&C was significantly higher than for those that were not ($Mdn = 3$ vs. 2.67), $Z = 19$, $p < .001$. It then occurred to us that certain TEQ items might cue memories more frequently than others, so we counted, across participants, the number of times each item was marked C&C. The first few items had been more frequently marked C&C, and a significant inverse correlation between TEQ item number and C&C frequency suggested that this could have been due to fatigue effects ($r = -.546$, $N = 16$, $p = .029$). Hence, as a more conservative test of our theory, we ran a related samples Wilcoxon signed rank test on TEQ scores associated with only the second half of the items, and still found a significant effect ($Mdn = 3$ vs. 2.67), $Z = 26.5$, $p = .032$. 
11 Discussion: Experiment 3

As predicted, participants’ TEQ scores were higher on items that cued memories of specific occasions on which they behaved empathetically. This finding connects with research on response biases in self-report measures of socially desirable traits, which are known to be prone to self-serving response biases (e.g., Edwards, 1953; Ludeke, Weisberg, & DeYoung, 2013; but see Paunonen & LeBel, 2012). For example, our results dovetail nicely with a study showing that people who score highly on a questionnaire measuring the vividness of visual imagery also score highly on a measure of the tendency to respond in a socially desirable way (Allbutt, Ling, Rowley, & Shafiullah, 2011). Our work may also shed new light on that study. Most readers will agree that empathy, which we measured with the TEQ, is a socially desirable trait. The fact that our participants rated themselves as more empathetic when they had vivid memories of themselves acting empathetically raises the question of how exactly episodic representation interfaces with social desirability. Perhaps people are less willing to indulge a self-serving bias if they cannot think of concrete evidence corroborating it. And since visual imagery is a component of both remembered and imagined events (Mullally & Maguire, 2013), do these two different types of episodic representation influence socially desirable responding in the same way?

Also, a recent study found that people who scored highly on impression management scales (which researchers often use for screening purposes and to account for participants with a self-serving response bias) were actually more honest than low-scorers in that they were less likely to cheat in a task where they were able to win money (Zettler, Hilbig, Moshagen, & de Vries, 2015). An interesting question that our MemFUS approach could address in this area is whether high scores on impression management scales correlate with the differences in ratings between items (on other self-report measures) that are associated with “congruent” episodic memories and those that are not. If they did, it would suggest that people who are concerned with how they come across try to manage their impression “honestly”, that is, by basing it on concrete evidence from episodic memory (although, of course, this would not avoid the issue that people’s memories themselves may be subject to distortion by self-serving biases, among other things).

Another interesting angle to our finding is that it raises the question of how much other, more widely-used, self-report instruments (e.g., the Big Five trait taxonomy; John & Srivastava, 1999) rely on the rater having a healthy and vivid episodic memory. Future research could adapt our
MemFUS questionnaire to other such self-report measures. Before that is undertaken, however, the MemFUS would need to be formally validated in a larger sample. It could probably also use some revision and expansion; e.g., distinguishing between specific episodic memories, “general” past or future events (see Addis, Cheng, Roberts, & Schacter, 2011), and semantic memory (including self-conceptions of one’s own personality).

Finally, our finding also has implications for the administration of self-report measures in clinical populations that have been associated with memory deficits, including people with major depression (e.g., Howe & Malone, 2011; Söderlund et al., 2014), post-traumatic stress disorder (e.g., Brown et al., 2014), and epilepsy (e.g., St-Laurent et al., 2009, 2011, 2014). When collecting self-report data from these individuals (e.g., on their quality of life; Aydemir, Ozkara, Canbeyli, & Tekcan, 2004; Giovagnoli & Avanzini, 2000), clinicians may find it informative to use the MemFUS to help identify issues that require further follow-up or to weight certain responses differently for assessment purposes.

12 Review of Experiments 1–3

See Table 3 below for a comparative summary of results from Experiments 1 and 2.

<table>
<thead>
<tr>
<th>Result</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willingness to help</td>
<td>imagine self &gt; other</td>
<td>imagine self &gt; crossword</td>
</tr>
<tr>
<td></td>
<td>imagine self &gt; math</td>
<td></td>
</tr>
<tr>
<td></td>
<td>imagine other &gt; math</td>
<td></td>
</tr>
<tr>
<td>Subjective vividness</td>
<td>imagine self &gt; math</td>
<td>imagine self &gt; crossword</td>
</tr>
<tr>
<td></td>
<td>imagine other &gt; math</td>
<td></td>
</tr>
<tr>
<td>Internal detail</td>
<td>NA</td>
<td>imagine self &gt; crossword</td>
</tr>
<tr>
<td>Correlation:</td>
<td>All conditions</td>
<td>crossword only</td>
</tr>
<tr>
<td>subjective vividness ~ willingness to help</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation:</td>
<td>NA</td>
<td>Place detail in imagine self</td>
</tr>
<tr>
<td>internal detail ~ willingness to help</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In Experiment 1, we found that participants were more willing to help individuals in need if, after reading vignettes describing the plight of those individuals, they imagined themselves helping them instead of solving math problems. This was also true when they imagined someone else helping. We have termed this advantage of imagination over unrelated cognitive activity the “prosocial construction effect” (PCE), though it is a replication of previous work (Gaesser & Schacter, 2014). Further, imagining oneself helping led to greater willingness to help compared to imagining someone else; although the $p$-value was only marginally significant, the effect size was identical to that reported by Gaesser et al. (in press). This result indicates that, as we predicted, self-referential processing is an important component of the PCE, though not a necessary one. We also found that the more vividly participants reported imagining, the more willing they were to help. This was true even in the math condition, where participants were not asked to imagine helping, and where mean vividness ratings were significantly lower than in the imagine conditions. This suggests that, no matter how brief or spontaneous an episodic representation is, if vivid, it can have important cognitive effects. Finally, we found that self-reported trait empathy (TEQ scores) correlated with willingness to help in all conditions and with vividness ratings in the imagine conditions. Thus, intuitively, people with higher trait empathy are more willing to help others; they also have a more vivid imagination.

The most important finding from Experiment 1 was that self-referential processing contributes to the PCE. An informal examination of other subjective ratings collected in Experiment 1 suggested that the role of self-referential processing is not to increase vividness nor confidence in one’s own ability to help, but is related to feeling more compassionate.

### Table 3. Comparative summary of results from Experiments 1 and 2.

<table>
<thead>
<tr>
<th>Correlation:</th>
<th>NA</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>internal detail ~ subjective vividness</td>
<td>All conditions</td>
<td>No (trending)</td>
</tr>
<tr>
<td>Correlation:</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>TEQ ~ willingness to help</td>
<td>imagine self and other only</td>
<td></td>
</tr>
<tr>
<td>Correlation:</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>TEQ ~ subjective vividness</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Comparative summary of results from Experiments 1 and 2.
In Experiment 2, we found that imagining oneself helping leads to greater willingness to help compared to a crossword activity in which participants generated personal memories and other mental imagery unrelated to helping. This shows that the PCE is not due to episodic representation *per se*, essentially replicating previous work (Gaesser & Schacter, 2014). We also recorded participants’ descriptions of what they imagined and scored the transcripts for episodic (internal) detail content using an adapted version of the Autobiographical Interview (AI; Levine et al., 2002). Transcripts from the *imagine self* condition were richer in episodic detail than those from the *crossword* condition, and episodic “place” details in particular predicted greater willingness to help on *imagine self* trials. Not only does this corroborate previous work (Gaesser & Schacter, 2014) using a novel measure, but it shows it is episodic detail in particular, and not detailed thinking in general, that is related to helping. It also shows that the PCE can be demonstrated with a standardized measure of vividness and not only with subjective ratings. Surprisingly, in the *imagine self* condition, we found no correlation between subjective vividness ratings and helping, nor between vividness ratings and episodic detail in transcripts. We believe this was due to a lack of variance in vividness ratings in the *imagine self* condition; describing their imagined events in detail before rating vividness likely strengthened participants’ memory for their imagined events, so that they recalled all imagined events as being highly vivid when making their ratings. Finally, after administering the TEQ at the beginning of the experiment instead of at the end, we found no significant correlations between TEQ scores and helping. Although this appears to conflict with our Experiment 1 results, visual inspection of scatterplots and removal of an outlier suggests there was actually a trending relationship. Regardless, TEQ scores were not significantly lower in Experiment 2, indicating that our participants’ self-reported trait empathy was not inflated after imagining helping in Experiment 1.

The most important finding from Experiment 2 is the relationship between willingness to help and the proportion of episodic place details in transcripts of imagined helping scenarios. We had predicted that other types of episodic detail would also be related to helping; the lack of such results may be due to demand characteristics (i.e., participants providing an overabundance of verbal detail), and the sole place detail finding fits with other evidence that establishing spatial context plays a foundational role in episodic representation (Mullally & Maguire, 2013).

In Experiment 3, we used a novel “memory follow-up survey” (MemFUS) with the TEQ and found that participants rated themselves as more empathetic on TEQ items that cued specific
memories of their own empathetic behaviour. We discussed the implications of this finding for other self-report instruments, especially those that measure socially desirable traits or that are used in clinical settings.

13 Limitations and Future Directions

Experiment 1 provided modest evidence that self-referential processing is an important component of the PCE. It was limited, however, in its ability to characterize the precise role that self-reference plays. Although we established that self-reference does not increase willingness to help by boosting subjective vividness or confidence in one’s own ability to help, the evidence that it increases feelings of compassion is tenuous; we simply observed that imagine self was associated with greater feelings of compassion compared to math while there was no such difference between imagine other and math. Future work, perhaps using a different paradigm, should investigate with more precision what the key contribution of self-reference is to the PCE.

A second limitation of Experiment 1, shared with Experiment 2, is that it was unable to infer a causal relationship between vividness and willingness to help. To address this issue, future research should investigate the PCE in people who have damage to the hippocampus/MTL or who have undergone surgical excisions in this brain area. Others have found that such individuals are impaired in both remembering (e.g., Rosenbaum et al., 2005) and imagining events (e.g., Hassabis et al., 2007; Kwan, Carson, Addis, & Rosenbaum, 2010; Race et al., 2011). Particularly interesting are studies in people who have temporal lobe epilepsy (TLE) or who have undergone anterior temporal lobectomy (ATL) surgery for relief of epileptic seizures. While these patients can recall the general narrative details of their past experiences, their memories are less vivid compared to those of healthy individuals due to impaired recollection of specific perceptual and temporal details (St-Laurent et al., 2014; St-Laurent, Moscovitch, Levine, & McAndrews, 2009; St-Laurent, Moscovitch, Tau, & McAndrews, 2011). The ability to imagine future events per se has not been specifically tested in TLE, although one study found that when TLE patients are asked to generate solutions to open-ended problems, their solutions contain a scant number of episodic details compared to solutions generated by healthy young adults (Sheldon et al., 2011). Also, as mentioned above, there is ample evidence that people who have very similar brain damage to that found in TLE/ATL patients, albeit of a different etiology, have difficulty both remembering and imagining events (for review see Mullally & Maguire,
If the PCE is driven by the vividness or episodic detail content of imagined helping events, then individuals with hippocampal damage or excisions should not exhibit the same increase in willingness to help that healthy individuals do after performing this task, because they are unable to imagine very detailed events. Such work may shed light on other findings that people without a healthy hippocampus are less affected by empathy induction exercises (Beadle, Tranel, Cohen, & Duff, 2013) and have smaller social networks (Davidson, Drouin, Kwan, Moscovitch, & Rosenbaum, 2012; see also Stiller & Dunbar, 2007) compared to other people.

Perhaps the greatest limitation of Experiment 3 is that it did not distinguish the influence of episodic vs. semantic memory on self-report ratings. The comparison we drew with our MemFUS approach was between TEQ items associated with episodic memories and all other items, which included items associated with semantic memories. It would be interesting to find out whether the effect of episodic memory on self-report is different from that of semantic memory, or from that of imagination, for that matter. Research on “temporal discounting” has shown that when people are given the choice between a small, immediate reward and a large, delayed reward, they are more likely to opt for the delayed reward when they vividly imagine using it in a specific future event, compared to when they use their general knowledge to conceive of ways to use it (Benoit, Gilbert, & Burgess, 2011). Further, potential differences should be investigated not only from a quantitative point of view (e.g., does one type of thinking lead to a “bigger” response than another?), but also from a qualitative one. For example, it could be that certain types of responding are enhanced by episodic memory retrieval and others by semantic memory retrieval, but both types of responding are enhanced to the same degree.

Finally, moving beyond questionnaires and prosocial cognition, future research on the link between episodic representation and empathy should make use of paradigms that are sensitive to real-world behaviour in both healthy and clinical populations.
References


Appendices

Appendix A – Examples of Vignettes and Math Problems

Vignettes

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>This person just finished eating at a restaurant and is feeling ill.</td>
</tr>
<tr>
<td>2</td>
<td>This flower delivery person, who usually uses their own car for deliveries, finds out their car was stolen the day before Valentine’s Day.</td>
</tr>
</tbody>
</table>

Math Problems

<table>
<thead>
<tr>
<th></th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Carlos found 50 seashells on a beach, and he gave some to Sophia. He has 42 seashells left. How many did he give to Sophia?</td>
</tr>
<tr>
<td>2</td>
<td>After eating at a restaurant, Emily, Kate, and Isaac decided to divide the bill evenly. If each person paid 32 dollars, what was the total of the bill?</td>
</tr>
</tbody>
</table>
Appendix B – MemFUS

Memory Follow-Up Survey (MemFUS), page 1

1. “When someone else is feeling excited, I tend to get excited too”
   Did you think about a memory from your own life when you were making a rating for the above statement on the survey you were filling out earlier? If you did not, that is okay. You do not have to think of one now; you can simply skip to the next page of the survey.

☐ Yes ☐ No

If you thought about a memory while filling out the survey earlier:
   In the event you remembered, how were you acting?

☐ I acted in a way that FITS with the statement

☐ I acted in a way that does NOT fit with the statement.

In the box below, briefly describe the memory you had when you were filling out the survey earlier. If you did NOT think of one at that time, leave it blank.

<table>
<thead>
<tr>
<th>Please write a brief description of what happened in this memory:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

When you thought about this memory (when you were filling out the survey) it was:

<table>
<thead>
<tr>
<th>1 . . . . . . . . . 2 . . . . . . . . . 3 . . . . . . . . . 4 . . . . . . . . . 5 . . . . . . . . . 6 . . . . . . . . . 7</th>
<th>Vague</th>
<th>Detailed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 . . . . . . . . . 2 . . . . . . . . . 3 . . . . . . . . . 4 . . . . . . . . . 5 . . . . . . . . . 6 . . . . . . . . . 7</td>
<td>Fragmented</td>
<td>Coherent</td>
</tr>
</tbody>
</table>