Seeing Scary: Predicting Variation in the Scariness of the Mental Representations of Spiders

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Alison Isobel Young

Graduate Program in Psychology

The Ohio State University

2014

Dissertation Committee:

Russell H. Fazio, Advisor

Richard E. Petty

Duane T. Wegener
Abstract

The current studies apply reverse correlation image classification (RCIC) techniques to estimate the mental representation of a spider among participants who vary in their fear of spiders. This is the first attempt that we know of to use RCIC techniques to both estimate individuals’ mental representations of a complex, non-facial object and to test whether those representations vary in a meaningful way. We find evidence in a pilot study that the RCIC technique is adaptable to non-facial, but complex, stimuli (spiders). In Study 1, we find that the mental representation of a typical spider among participants higher in fear of spiders is rated by objective judges as looking scarier and more threatening than the mental representation among participants lower in fear of spiders. In Study 2, we find evidence of this effect for a specific spider image to which participants were exposed earlier in the study. Methodologically, our findings suggest that RCIC techniques can be used to understand individual differences in the representations of complex, non-facial stimuli. At a more theoretical level, the findings illustrate how perceptions can be influenced by the emotions and evaluations that individuals associate with an object.
Dedication

Dedicated to my parents
Acknowledgments

My thanks to my advisor, Russell Fazio, who has been extremely helpful to me during the preparation of the research reported here and throughout my time at The Ohio State University. I would also like to acknowledge my fellow graduate students in the Fazio lab (Eva Pietri, Matt Rocklage, Pete Zunick, Elise Bui and Aaron Hatchett) who provided valuable feedback on this work, as well as the Social Psychology program at Ohio State, which has prepared me well for a career as a social psychologist.

I would like to thank my family for all of their support throughout the years. I love you all very much. Finally, I would like to thank God for His constant provision and grace.
Vita

2003.................................................................Fillmore Central School

2007.................................................................B.A. Psychology, Houghton College

2010.................................................................M.A. Psychology, The Ohio State University

2010 to present ..............................................Graduate Teaching Associate, Department of Psychology, The Ohio State University

Publications


Fields of Study

Major Field: Psychology
Table of Contents

Abstract ........................................................................................................................................ ii

Acknowledgments .................................................................................................................. iv

Vita ........................................................................................................................................... v

Table of Contents ................................................................................................................... vii

List of Tables ........................................................................................................................ x

List of Figures ....................................................................................................................... xi

Chapter 1: Introduction ......................................................................................................... 1

Biased Construal and Spider Fear ....................................................................................... 2

Assessing the Mental Representation ............................................................................... 4

RCIC and Biased Construals .............................................................................................. 7

The Current Research ....................................................................................................... 8

Chapter 2: Pilot Study ......................................................................................................... 10

Method .............................................................................................................................. 10

Participants ......................................................................................................................... 11

Stimuli for Image-Classification Task ............................................................................. 11
List of Tables

Table 1. HLM Regression Coefficients for Study 1: Predicting the Scariness of Participants' Weighted Typical Spider Classification Images (CIs). ............................... 27

Table 2. HLM Regression Coefficients for Study 2: Predicting the Scariness of Participants' Over-Weighted Specific Spider Classification Images (CIs). ......................... 39
List of Figures

Figure 1. An example of the spider stimuli used in the image classification task. The base image from which stimuli were derived is depicted at the top; sample stimuli are below. Participants selected which of the two spiders looked scarier for 450 pairs. .................... 13

Figure 2. Dichotomous (top row) and weighted (bottom row) CI pairs for the pilot study. The CIs based on trials in which the spider looked scarier to participants are on the right. The CIs based on non-selected trials are on the left. ........................................................ 17

Figure 3. Dichotomous (top row) and weighted (bottom row) CI pairs for Study 1. The CIs based on participants lower in Fear of Spiders are on the left. The CIs based on participants higher in Fear of Spiders are on the right. ..................................................... 25

Figure 4. Weighted (bottom row) CI pair for Study 2. The CI based on participants lower in Fear of Spiders is on the left. The CI based on participants higher in Fear of Spiders is on the right. ....................................................................................................................... 37

Figure 5. The 105419 pixels included in participant CI – pilot ‘scary’ CI correlations in Study 2. ............................................................................................................................................ 40
Chapter 1: Introduction

The idea that we see the world in a biased fashion is not new. Jerome Bruner argued more than half a century ago that perception itself is influenced by top-down categorization processes in the perceiver (Bruner, 1954). Whether while walking down the sidewalk I mistakenly perceive a garden hose as a snake depends on a number of variables that make me ready to perceive a snake (e.g., snakes have been recently primed by my watching a special on Animal Planet, I have learned that the neighbor’s daughter lost her pet snake yesterday, etc.). Much work in the domain of social cognition has focused on these biased construals. For example, Bruner & Goodman (1947) found that poorer ten-year-olds, whose subjective need for money was high, overestimated the size of coins relative to richer ten-year olds. Hastorf & Cantril (1954) found evidence that participants’ team affiliation biased whether they deemed a particular football play an infraction or not. Fazio, Roskos-Ewoldsen, & Powell (1994) found that participants asked to serve as line judges in a computerized tennis game were more likely to (that is, perceptually ready to) judge a tennis ball as “in” if it was hit by a liked versus disliked confederate. These results occurred in spite of the fact that participants were told that the performance of the confederate would be objectively rated by the computer, and that the participants’ own ratings were to provide information about the screen resolution only.
Similar biasing effects have been found on geographical slant perception (e.g., Bhalla & Proffitt, 2000; Schnall, Harber, Stefanucci, & Proffitt, 2008; Riener, Stefanucci, Proffitt, & Clore, 2011), perception of object size (e.g., Witt & Proffitt, 2005; Linkenauger, Witt, & Proffitt, 2011; van Ulzen, Semin, Oudejans & Beek, 2008), and distance (e.g., Witt, Linkenauger, Bakdash, Augustyn, Cook & Proffitt, 2008; Lessard, Linkenauger, & Proffitt, 2009; Balcetis & Dunning, 2010; Cole, Balcetis & Zhang, 2013). In all cases, subjective need, ability, or desire impacted perception.

Similar work in biased construals has focused on an alternative influence: fear. Slopes seem steeper to individuals who are made to feel afraid relative to unafraid controls (Stefanucci, Proffitt, Clore, & Parekh, 2008). Distances from threatening individuals seem closer than from neutral individuals (Cole, Balcetis, & Dunning, 2013). Heights, when viewed from above, seem higher for those higher in state and trait fear of heights (e.g., Teachman, Stefanucci, Clerkin, Cody & Proffitt, 2008; Stefanucci & Proffitt, 2009; Stefanucci & Storbeck, 2009) as well as for conditions in which the outcome of falling is particularly dire (Stefanucci, Gagnon, Tompkins & Bullock, 2012). Fear seems to be related to perceptual distortions of stimuli relevant to that fear.

**Biased Construal and Spider Fear**

Work in the clinical domain has found some evidence that a specific fear, spider fear, impacts perception. Spiders, for instance, draw attention more rapidly in fearful individuals relative to less fearful individuals (Vrijsen, Fleurkens, Nieuwboer & Rinck, 2009; Miltner, Krieschel, Hecht, Trippe & Weiss, 2004; Rinck, Reinecke, Ellwart, Heuer & Becker, 2005; Van Strien, Franken & Huijding, 2009; Peira, Golkar, Larsson & Wiens,
In a binocular rivalry paradigm, Gerdes & Alpers (2014) found evidence that these attentional effects are especially pronounced when active avoidance is not possible.

Not only do spiders draw attention for fearful individuals, but there is also some evidence that spiders seem to be construed differently. Clinical work has suggested that self-reported spider imagery for phobic individuals is significantly higher in clarity, vividness, anger, and movement – particularly, movement toward the phobic individual – than for non-phobic individuals (e.g., Hekmat, 1987; Weerts & Lang, 1978; Riskind, Moore & Bowley, 1995). Researchers have found similar results for perceived spiders: a photograph of a tarantula appears to individuals higher in spider fear to be moving toward them faster relative to individuals lower in spider fear (Riskind, Kelly, Moore, Harman & Gaines, 1992); a spider in a bowl is rated by participants higher in fear of spiders to be jumping more frequently, as moving towards them, and as less predictable relative to participants lower in spider fear (Rachman & Cuk, 1992); projected spiders moving towards participants across a tabletop appear to be moving faster relative to non-threatening objects (Witt & Sugovic, 2013; note that this study did not moderate speed perceptions by spider fear). Participants who feel more threatened by a tarantula estimate that tarantula as being closer to them (Cole, Balcetis & Dunning, 2013) and in a clinical sample, participants who experience higher state anxiety in the presence of a spider estimate that spider as larger in size relative to participants who experience less state anxiety (Vasey, Vilensky, Heath, Harbaugh, Buffington & Fazio, 2012).

Thus, there is an abundance of evidence that spiders are subjectively judged to have more threatening characteristics (e.g., they look larger, faster, and more
unpredictable) by individuals with higher levels of spider fear. Based on this evidence, might it also be the case that spiders are mentally represented differently in the minds of people who are afraid versus unafraid of them? In other words, could people higher in spider fear picture spiders in their “mind’s eye” as looking scarier and more threatening relative to people lower in spider fear? Does their mental representation of a prototypical spider or even a specific spider to which they are directly exposed actually appear more threatening?

Assessing the Mental Representation

In order to test whether a spider in the “mind’s eye” of a fearful person looks objectively scarier than that in the mind of a less fearful person, we employed reverse correlation image classification (RCIC) techniques. These techniques were first employed in auditory research (Ahumada & Lovell, 1971) and in visual psychophysics research (Ahumada, 1996; Beard & Ahumada, 1998; see Murray, 2011 for a review), and have been applied to research on feature detection (Neri & Heeger, 2002), processing of upright vs. inverted faces (Sekuler, Gaspar, Gold, & Bennett, 2004), illusory contours (Gold, Murray, Bennett, & Sekuler, 2000) and letter discrimination (Watson, 1998). In RCIC, participants are presented with stimuli that vary randomly because they consist of a single invariant base image (e.g., a face, an illusory square, two lines, etc.) superimposed with randomly varying noise, which serves to change the look of the base image. Participants are asked to classify these stimuli according to a particular prompt (e.g., which of these two faces do you see, which of these two illusory squares do you see, are these two lines aligned, etc.). From participants’ responses, researchers can
generate a classification image (CI), an average of selected noise patterns across all trials, which indicates what features of the randomly-varying stimuli predict participants’ classifications. In a sense, the CI correlates participants’ responses with variations in the randomly-varying noise (Eckstein & Ahumada, 2002).

In theory, RCIC allows researchers to assess internal representations of various objects in the world. It has been successfully employed to generate estimates of these representations even when no base image is included in the stimuli. Gosselin, Bacon, & Mamassian (2004), for example, presented two participants with 20,000 images which were essentially pure, random noise (specifically, random-dot stereograms). Across these trials, they were told to report whether they detected the presence of a large ‘+’ sign despite the fact that no ‘+’ sign was ever present in the images. Summing all responses, the authors demonstrated that participants were able to ‘pull out’ a ‘+’ sign in random noise, providing an estimate of their internal representation of the symbol. Gosselin & Schyns (2003) reported similar findings with the letter ‘S.’

Mangini & Biederman (2004) adapted the technique to the estimation of faces by using randomly-varying sinusoidal noise rather than Gaussian noise. In three separate tasks, these researchers superimposed the randomly-varying sinusoidal noise on a particular base face (e.g., a morph of John Travolta and Tom Cruise), which had the effect of randomly changing the features of that face. The inclusion of the base face rather than noise alone served to reduce the number of trials by constraining noise variation to facial features, rather than the presence or absence of any face. Across 390 trials, participants were asked to classify each of 390 faces with noise added according to
a given prompt (e.g., classify whether the image looked more like John Travolta or Tom Cruise). Since the base face itself did not change from trial to trial, the only information participants were using to make their decisions was the superimposed noise.

Based on participants’ classifications, Mangini & Biederman were able to generate classification images which provided an estimate of whatever participants were imagining while completing each task, whether they were imagining the face of a particular person (e.g., John Travolta or Tom Cruise), a particular emotion (e.g., happy or unhappy; see also Kontsevich & Tyler, 2004) or a particular gender (male or female). In other words, reverse correlation image classification when applied to face stimuli allowed these researchers to get a glimpse of the features an individual participant associated with a given face. According to Mangini & Biederman (2004), one advantage of this method is that the freely-varying features are not limited to a select few the experimenter may be interested in. Another advantage is that the stimuli are unbiased by the questions the experimenter is interested in testing. Since the noise varies randomly, it is the participant’s choices, not the experimenter’s, that impact what features contribute to that participant’s mental representation of a given face.

Following Mangini & Biederman’s (2004) technique, many other studies in social psychology used reverse correlation techniques to understand the mental representations of faces. Researchers have used the technique to assess more general mental representations of prototypical trustworthiness and dominance (Dotsch & Todorov, 2012), representations of the warmth and competence dimensions in prototypical social groups (Imhoff, Woelki, Hanke, & Dotsch, 2013), representations of ingroup versus

**RCIC and Biased Construals**

A fair amount of recent work has leveraged RCIC to assess biased construals; that is, how the mental representation of objects might vary depending on participants’ attitudes, stereotypes, cultural backgrounds, etc. Some researchers have explored the biased representation of prototypical faces. For example, Dotsch, Wigboldus, Langner & van Knippenberg (2008) demonstrated that participants who scored more highly on an implicit measure of prejudice against Moroccan people represented a prototypical Moroccan face in a way that objective judges viewed as more criminal and less trustworthy. Imhoff, Dotsch, Bianchi, Banse & Wigboldus (2011) found that German and Portuguese participants tended to mentally represent a typical European face in a way that was more typical of their own national identity. Jack, Caldara & Schyns (2012) found evidence that Western Caucasian and East Asian participants represented facial expressions of basic emotions in culturally specific ways; that is, Western Caucasian participants tended to feature the eyebrow and mouth regions in their representations of six basic emotions, whereas East Asian participants featured the eye regions. Dotsch, Wigboldus, & van Knippenberg (2013) actually manipulated the criminality and trustworthiness of the behaviors of a group of people. Participants were presented with faces of exemplar group members along with their behaviors. The researchers then assessed the face participants expected typical group members to have (a prototypical
group face) using RCIC techniques. They found that when the group members’ behaviors were suggestive of high criminality and low trustworthiness, this prototypical group face was judged by objective raters as looking more criminal and less trustworthy than when the group members’ presumed behaviors were indicative of low criminality and high trustworthiness.

Still others have found evidence that even the way in which participants mentally represent specific individuals might be biased. Karremans, Dotsch & Corneille (2011) found, for instance, that an attractive male, whose photograph was presented for one minute at the beginning of the study, was mentally represented less attractively among participants who were romantically involved versus romantically uninvolved. Recent research indicates that these biasing effects do not seem to be limited to novel stimuli seen for short amounts of time but occur even for very familiar faces. Young, Ratner & Fazio (2013) found that during the 2012 election in the swing-state of Ohio, Mitt Romney, a well-known presidential candidate, was mentally represented in a more trustworthy way among Republican participants relative to Democrat participants.

The Current Research

Our work has three main goals. The first is methodological. Past RCIC work has been applied to low-level perceptual stimuli (e.g., lines and contours) in the psychophysics realm. It has been applied to higher-level, more complex stimuli – faces – in the social psychological domain. However, RCIC has yet to be applied to non-facial stimuli with high-level features. That RCIC has been used successfully with face stimuli despite their complexity is reasonable considering most adults are expert face processors...
(Diamond & Carey, 1986). Whether RCIC can be used for other high-level stimuli for which most adults don’t have as much expertise, like spiders, remains to be seen.

Our second goal, assuming the successful implementation of RCIC for spider stimuli, is to test whether individuals’ mental representations of spiders vary in a meaningful way. As we detailed earlier, biased construal is an established finding across multiple domains. Prior research has suggested that fear-related objects are construed in a biased way, particularly by people high in a given fear. Much work has been done on this in the realm of spider phobia, with judgments as noted earlier of a spider’s size, speed, proximity, etc. But no work has yet focused on how the gestalt visual representation of a spider might be biased in accord with spider fear.

Our third goal is to test whether the representations of spiders vary for both a prototypical spider and a specific spider exemplar. Past work on RCIC with face representations has focused on these two object types separately; we sought to test whether we could find similar biasing effects in both cases.

In all three of the following studies, we based our methods on those used by Dotsch et al. (2008). Each study consisted of two phases. In the first phase, participants completed an image-classification task (Dotsch et al., 2008; Mangini & Biederman, 2004). We additionally modified the paradigm to incorporate a confidence scale, allowing us to weight the ultimate classification images (CIs). In the second phase, objective raters assessed the CI + base image composites.
Chapter 2: Pilot Study

Before we could use RCIC techniques to assess meaningful variation in the mental representations of spiders in people higher versus lower in fear of spiders, we first had to test whether participants could generate spiders that looked objectively scary during the image classification task. Is it possible for participants to, in effect, create a scary image from a base image of a spider and random noise? In our pilot study, we asked a small group of participants to choose which of two randomly varying spider images looked scarier and averaged their selections together to see whether the chosen spider CI looked scarier than the non-chosen spider CI. If so, we would be able to use the technique to assess future participants’ mental representations of a typical spider, or of a specific spider they had seen earlier.

Method

The pilot study consisted of two phases. In phase one, participants completed the image classification task and in Phase 2, objective raters assessed the classification images (technically, the CIs superimposed over the base image) generated by participants in Phase 1. The goal of the pilot was to test whether RCIC techniques could be used to ‘pull out’ a scary-looking spider.
Participants

Phase 1 participants were 13 undergraduates who took part for course credit. Phase 2 participants \((N = 84; \text{45 men, 39 women; age range 18-65, mean age } = 32.64)\) were a sample of adults from Amazon’s Mechanical Turk (MTurk) Web site who rated pairs of CIs based on chosen (scary) versus non-chosen (non-scary) trials. These participants were naïve to how the spider images were generated and were paid 75 cents for 8 minutes of their time.

Stimuli for Image-Classification Task

All stimuli for this task were generated from the same base image of a jumping spider. After exploring other spider types, we selected this spider because its body was substantial enough not to be obscured beyond recognition by superimposed noise while also allowing the noise to change its appearance. The spider image was blurred slightly using a Gaussian filter \((\sigma = 4 \text{ pixels})\), but many details, such as the hairs on its legs and the dark lines along its body, were still visible (see Figure 1). The image was converted to grayscale. To change the appearance of the spider on each trial, we superimposed noise patterns on the 512 x 512 image (see Fig. 1). These noise patterns consisted of 4,092 superimposed truncated sinusoid patches, each of which spanned two cycles of a sine wave. The patches were summed across six orientations \((0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, \text{and } 150^\circ)\), two phases \((0 \text{ and } \pi/2)\), and five spatial frequencies \((1, 2, 4, 8, \text{and } 16 \text{ patches per image})\), with random contrasts (amplitudes) to create each of the 450 noise patterns we used. For a more complete description of patch creation, see Mangini & Biederman,
2004. All image pairs (base + noise pattern or base + inverse noise pattern) were then scaled to 256 X 256 pixels prior to presentation.
Figure 1. An example of the spider stimuli used in the image classification task. The base image from which stimuli were derived is depicted at the top (Collins, 2011); sample stimuli are below. Participants selected which of the two spiders looked scarier for 450 pairs.

**Procedure: Phase 1**

*Image-classification task.* Participants were told that they would see a series of spiders in pairs and that the images would be blurry and hard to see because we added
noise to them. They were instructed to select, for each of 450 image pairs, which spider they believed looked scarier. They responded on a six-point scale to indicate whether they thought the spider on the left or the right was much scarier, scarier, or slightly scarier. The question (“Which spider looks scarier?”) and the scale labels were on screen for each of the 450 trials. After completing the image-classification task, Phase 1 participants were thanked for their time and debriefed.

**Generating a “scary” spider.** To generate an estimate of participants’ collective idea of a scary-looking spider, we used reverse-correlation data-reduction techniques. We averaged the parameters which defined the noise patterns which participants had selected as looking scarier across all 450 trials per participant, resulting in 4092 mean parameters per participant. We then averaged these 4092 mean parameters again across participants and superimposed the noise pattern based on these averaged parameters over the base spider image to generate a classification image (CIs). This CI represented participants’ overall idea of a scary-looking spider. We also similarly averaged the parameters of the non-selected noise patterns across participants to get a CI which would represent a non-scary-looking spider.

Because Phase 1 participants responded using a 6-point scale, we were able to generate two types of CIs. The first was a dichotomous CI. For this CI, we recoded participant responses as a purely dichotomous choice regardless of the extremity of their key press. This CI type was analogous to earlier reverse correlation work in which participants made a simple binary choice between the two images on each trial.
The second type of CI we generated was a weighted CI. In this case, trials for which participants responded that one image was “much scarier” than the other were included three times in the calculation of the CI, trials for which participants responded that one image was “scarier” than the other were included twice in the CI, and trials for which participants responded that one image was “slightly scarier” than the other were included once in the CI. The weighted CI thus leveraged the fact that on certain trials, one spider looked more obviously scarier than the other to the Phase 1 participant.

**Procedure: Phase 2**

In this phase, independent raters on MTurk compared the “scary” CI to the “non-scary” CI (dichotomous and weighted) on a series of variables. The two CIs were displayed side-by-side, one on the left and one on the right (counterbalanced across raters). Raters were asked to indicate in which of the two photographs the spider looked more threatening, frightening, and scarier (order randomized). They responded to these first three items on 6-point continua (anchors were *looks much more threatening/much more frightening/much scarier in the photo on the LEFT* and *looks much more threatening/much more frightening/much scarier in the photo on the RIGHT*). Raters then indicated in which image the spider made them feel more anxious, also on a 6-point continuum (anchors were *the spider in the LEFT image makes me feel much more anxious* and *the spider in the RIGHT image makes me feel much more anxious*).

After providing these ratings, raters also indicated the extent to which they were afraid of spiders on a 4-point scale (anchors were *I am not at all afraid of spiders* and *I*...
am extremely afraid of spiders). They then completed demographic items and were presented a debriefing statement.

Results

Phase 2 participants ranged in their responses to the fear of spiders item ($M = 2.58$ on a 4-point scale, $SD = .90$), but since this was not a significant predictor in the pilot, we will not discuss it further.

Dichotomous Classification Images

All four ratings of the dichotomous CI pair were recoded and centered around zero such that higher numbers indicated that the CI based on selected trials was rated as scarier, more threatening, etc. than the CI based on non-selected trials. Because all four ratings were highly interrelated, we averaged them to get a single mean rating per Phase 2 rater ($\alpha = .97$). Phase 2 raters indicated that the CI based on Phase 1 participants’ dichotomized choices looked significantly scarier, more threatening, more frightening, and more anxiety-inducing than the CI based on non-chosen trials ($M = .88$, $t(83) = 4.73$, $p < .0001$).

Weighted Classification Images

We found similar effects when we focused on the weighted CI image pair. Again, a composite of all four ratings was very reliable ($\alpha = .97$). As before, the spider CI based on chosen trials in Phase 1 was rated as scarier, more threatening, more frightening, and more anxiety-inducing than the spider CI based on non-chosen trials ($M = .88$, $t(83) = 4.63$, $p < .0001$). Figure 2 displays the chosen vs. non-chosen dichotomous and weighted CI pairs.
Figure 2. Dichotomous (top row) and weighted (bottom row) CI pairs for the pilot study. The CIs based on trials in which the spider looked scarier to participants are on the right. The CIs based on non-selected trials are on the left.
Pilot Discussion

The pilot study suggests that RCIC methods have the potential to differentiate between a scary versus a non-scary-looking spider. Since the $t$-ratios for the dichotomous and weighted CI pairs were nearly identical, it is unclear whether weighting the image average based on trials in which Phase 1 participants saw a clearer difference between the scariness of the two spiders allowed us to better detect differences between the scary vs. the non-scary images. However, we decided to continue using the scale paradigm in case weighted CIs might allow us to detect more subtle differences than dichotomous CIs.
Chapter 3: Study 1

The findings from our pilot study indicated that we were able to adapt RCIC techniques to generate spider CIs which differed in terms of scariness. Given this, we sought to apply this method to understand whether participants might mentally represent spiders in different ways based on their attitudes towards spiders. Specifically, we hypothesized that participants who are more afraid of spiders might represent those spiders as looking scarier and more threatening than participants who are less afraid of spiders. In our first test of this hypothesis, we focused on participants’ general representation of a typical spider.

Method

Participants

Phase 1 participants were 67 undergraduates (43 women, 24 men) who took part for course credit. Three Phase 1 participants were excluded, two for reporting that they did not take the image classification task seriously, and one for pressing the same key for the entire image-classification task. Our final Phase 1 N was 64 (43 women, 21 men). Two samples served as Phase 2 participants. The first was a sample of adults from Amazon’s Mechanical Turk (MTurk) Web site (N = 124; 69 men, 55 women; age range 18-65, mean age = 32.60) who rated pairs of CIs based on two groups of Phase 1
participants: those who were higher versus lower in Fear of Spiders. These raters were naïve to how the spider images were generated and were paid 50-75 cents for 5-8 minutes of their time. The second Phase 2 sample involved 91 additional MTurk participants (49 men, 42 women; age range 18-65; mean age = 34.30), also naïve to how the images were generated, who were paid $1.00 to rate each of the CIs generated by the 64 Phase 1 participants.

**Stimuli for Image-Classification Task**

All stimuli for this task were the same stimuli we used in the pilot study.

**Procedure: Phase 1**

**Image-classification task.** Participants were told that we were interested in their memory of objects they have encountered many times before. The specific memory object we asked them to focus on was a typical tarantula. In order to test their memory, we told them that they would see a series of images in pairs and that the images would be blurry and hard to see because we added noise to them. They were instructed to select, for each of 450 image pairs, which spider they believed more accurately depicted a typical tarantula. They responded on a six-point scale to indicate whether they were: very confident, confident, or would guess that the image on the left versus the right looked more like a typical tarantula. The question (“Which looks more like a typical tarantula?”) and the scale labels were on screen for each of the 450 trials.

**Fear of Spiders.** After the image-classification task, we assessed the extent to which Phase 1 participants were afraid of spiders using the Fear of Spiders scale (Szymanski & O’Donohue, 1995). The 18 items in this scale assess current avoidance of
and fear of harm from spiders (e.g., “If I came across a spider now, I would leave the room;” “Spiders are one of my worst fears”). All items employ a 7-point Likert scale from -3 (strongly disagree) to +3 (strongly agree). The scale had good reliability in this sample ($\alpha = .96$) and all 18 items were averaged together to create a single Fear of Spiders composite. Once participants had completed this scale, they responded to some demographic questions and then were debriefed.

Assessing mental representations of a typical tarantula. To generate an estimate of participants’ mental representation of a typical tarantula, we again used reverse-correlation data-reduction techniques. As with prior work using this method (Dotsch et al., 2008; Mangini & Biederman, 2004), we expected that on each trial, participants chose the spider image which best matched their mental image of a typical tarantula. We generated two types of estimates of these mental representations: participant-level and group-level. We generated participant-level estimates by averaging the parameters which defined each of the 450 noise patterns which were selected by each participant. This resulted in 4092 mean parameters per participant. We then superimposed noise patterns based on these mean parameters over the base spider image to create 65 separate classification images (CIs), one per participant. These CIs provided an estimate of each participant’s mental representation of a typical tarantula. We generated group-level estimates by further averaging the mean parameters for the subgroup of participants higher in Fear of Spiders as well as the subgroup of participants lower in Fear of Spiders. This resulted in a pair of CIs which provided an estimate of a
particular group’s mental representation of a typical tarantula (e.g., for those higher versus lower in Fear of Spiders).

Similar to the pilot study, because Phase 1 participants responded using a 6-point scale, we were able to generate both dichotomous and weighted CIs for the group-level estimates. For dichotomous CIs, we recoded participant responses as a purely dichotomous choice regardless of the extremity of their key press. For weighted CIs, trials for which participants responded that they were “very confident” that one image looked more like a typical tarantula were included three times in the CI; trials for which participants responded that they were “confident” were included twice in the CI, and trials for which participants responded that they would “guess” were included once in the CI. The weighted CI essentially weighted confident trials more heavily in the ultimate CI than trials on which participants guessed.

Procedure: Phase 2.

Ratings of group-level CIs. In this phase, a sample of 124 MTurk participants compared the typical tarantula CIs from the group of Phase 1 participants high versus low in fear of spiders (dichotomous and weighted) on a series of variables. These group-level CIs were displayed side-by-side, one on the left and one on the right (counterbalanced). Raters were asked to indicate in which of the two photographs the spider looked more threatening, frightening, and scarier (order randomized). They responded to these first three items on 6-point continua (anchors were looks much more threatening/much more frightening/scarier in the photo on the LEFT and looks much more threatening/much more frightening/scarier in the photo on the RIGHT). Raters then indicated in which
image the spider made them feel more anxious, also on a 6-point continuum (anchors were *the spider in the LEFT image makes me feel much more anxious* and *the spider in the RIGHT image makes me feel much more anxious*).

After providing these ratings, MTurk raters also indicated the extent to which they were afraid of spiders on a 4-point scale (anchors were *I am not at all afraid of spiders* and *I am extremely afraid of spiders*). They completed demographic items and were debriefed.

**Ratings of participant-level CIs.** A separate sample of 91 independent raters on MTurk rated the 64 weighted participant-level CIs generated in Phase 1. They were asked to view each CI and report how scary/threatening/frightening the tarantula looked on a 10-point scale (anchors were 0 – *not at all* and 9 – *extremely*). They were then given the full Fear of Spiders scale and debriefed.

**Results**

Phase 1 participants varied in terms of their Fear of Spiders (*M* = -.82, *SD* = 1.55). Phase 2 MTurk participants who rated each of the 64 Phase 1 CIs also ranged in Fear of Spiders (*M* = -.28, *SD* = 1.65). Phase 2 MTurk participants who rated the group-level CIs ranged in their responses to the single 4-point fear of spiders item (*M* = 2.52, *SD* = .87).

**Predicting the Scariness of Group-Level CIs**

MTurk participants rated the group-level CIs based on the group of Phase 1 participants higher in Fear of Spiders (upper 15 percent: *n* = 10; scale value 1.29 and above) and lower in Fear of Spiders (lower 15 percent: *n* = 9; scale value -2.51 and below). Responses were recoded and centered around zero such that higher numbers
indicated that the CI based on participants higher in Fear of Spiders was rated as scarier, more threatening, etc.

**Dichotomous classification image averages.** Because all four ratings of the dichotomous CI pair were highly related ($\alpha = .95$), we averaged these ratings to get a single mean rating per Phase 2 rater. As predicted, the dichotomous tarantula CI derived from Phase 1 participants higher in Fear of Spiders was rated as scarier, more threatening, more frightening, and more anxiety-inducing than the dichotomous tarantula CI based on Phase 1 participants lower in Fear of Spiders ($M = .32, t(123) = 2.18, p = .03$). Raters’ responses to the single fear of spiders item did not predict their ratings ($p = .63$).

**Weighted classification image averages.** When we weighted the group-level CIs by participants’ confidence in their selections, we again found the predicted effect. Again, all four ratings correlated strongly and were collapsed into a single item ($\alpha = .96$). The weighted tarantula CI derived from Phase 1 participants higher in Fear of Spiders was rated as scarier, more threatening, more frightening, and more anxiety-inducing than the weighted tarantula CI based on Phase 1 participants lower in Fear of Spiders ($M = .45, t(123) = 3.04, p = .003$). Note that this effect is somewhat larger than was the case for the dichotomous CI pair (dichotomous pair Cohen’s $d = .39, r = .19$; weighted pair Cohen’s $d = .55, r = .26$). Raters’ responses to the single fear of spiders item did not predict their ratings ($p = .92$). Figure 3 displays the dichotomous and weighted CI pairs.
Figure 3. Dichotomous (top row) and weighted (bottom row) CI pairs for Study 1. The CIs based on participants lower in Fear of Spiders are on the left. The CIs based on participants higher in Fear of Spiders are on the right.

Predicting the Scariness of Participant-Level CIs

Since the group-level results were strongest for the weighted CIs, the 91 Phase 2 raters rated weighted participant-level CIs. We ran a two-level HLM analysis involving...
5824 observations based on 64 Phase 1 participants nested in 91 Phase 2 raters. The mean rating across the 5824 observations was 5.71 and the standard deviation was 2.15. The model predicted the rating of how scary/threatening/frightening a given Phase 1 participant’s tarantula CI was from Phase 1 participants’ Fear of Spiders (entered group-mean centered) at level one and from Phase 2 participants’ Fear of Spiders (entered grand-mean centered) at level 2. We modeled all coefficients as fixed if the associated error term was not significantly different from zero, using a generous cut-off $p$-value of 0.2 as recommended by Nezlek (2011).

As predicted, we found a main effect of Phase 1 participant Fear of Spiders scores on the ratings of how scary/threatening/frightening their mental representation of a typical tarantula was ($\gamma = .02$, $t(5731) = 3.65$, $p < .001$). The higher in Fear of Spiders the Phase 1 participant was, the scarier, more threatening, and more frightening the weighted typical tarantula CI generated by that participant was judged to be by independent raters. We also found a main effect of rater Fear of Spiders such that the higher in Fear of Spiders the rater was, the scarier, in general, all 64 CIs were judged to be ($\gamma = .59$, $t(89) = 5.48$, $p < .001$). See Table 1 for the results for this model.
Table 1. HLM Regression Coefficients for Study 1: Predicting the Scariness of Participants' Weighted Typical Tarantula Classification Images (CIs).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>5.71</td>
<td>(0.17)</td>
<td>***</td>
</tr>
<tr>
<td>Phase 2 Rater FOS ($\gamma_{01}$)</td>
<td>0.59</td>
<td>(0.11)</td>
<td>***</td>
</tr>
<tr>
<td>Phase 1 Participant FOS ($\gamma_{10}$)</td>
<td>0.02</td>
<td>(0.01)</td>
<td>***</td>
</tr>
<tr>
<td>Cross-level interactions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 1 Part. FOS X Rater FOS ($\gamma_{11}$)</td>
<td>-0.00</td>
<td>(0.00)</td>
<td></td>
</tr>
</tbody>
</table>

Significance: + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed test).

Standard errors in parentheses.

**Discussion of Study 1**

These findings suggest that participants mentally represent a typical tarantula differently depending on the extent to which they are generally afraid of spiders. Participants who have higher fear of spiders mentally represent a tarantula in a way that is judged by objective observers as looking scarier and more threatening than a tarantula in the minds of participants lower in fear of spiders. Recall that pilot participants, when asked to choose which spider actually looked scarier, produced a CI which looked objectively scarier than a CI based on unselected spiders. Based on this, the fact that in Study 1, participants higher in fear of spiders produced a CI which also looked...
objectively scarier than the CI based on participants lower in fear of spiders suggests that high-fear participants seem to be responding in a similar way to the classification task – that is, basing their image classification selections at least somewhat on the scariness of the spiders. This, along with our objective rating data, suggests that the representation in memory of a typical tarantula looks scarier among high-fear versus low-fear participants.

However, these findings do not imply that fear of spiders necessarily affects how participants mentally represent a specific spider to which they have been exposed. Might participants’ fear of spiders more directly impact how they represent a spider they have just seen earlier in the same study? Study 2 sought to test this possibility.
Chapter 4: Study 2

In Study 2, participants were introduced to a specific spider in the context of many other image stimuli. They were then asked to choose during the image classification task which image looked more like the spider they had seen earlier. We hypothesized that even though participants had seen this specific spider image, their mental representations of that specific spider would be biased in line with their fear of spiders.

Method

Participants

Phase 1 participants were 99 undergraduates (51 women, 48 men) who took part for course credit. The data from seven Phase 1 participants were omitted because they did not appear to take the image classification task seriously (that is, more than half of their trials were completed at less than 200 milliseconds each). Our final Phase 1 N was 92 (49 women, 43 men). One sample of Phase 2 participants (N = 124; 69 men, 55 women; age range 18-70, mean age = 30.84) rated pairs of group-level CIs based on Phase 1 participants who were higher vs. lower in Fear of Spiders on Amazon’s Mechanical Turk (MTurk) Web site. These participants were naïve to how the spider images were generated and were paid between 50 and 75 cents for 5-8 minutes of their time. Another
sample of Phase 2 participants ($N = 90$; 47 men, 43 women; age range 19-60, mean age = 34.49) rated all 92 weighted CIs generated by the Phase 1 participants, and 74 additional Phase 2 undergraduate participants took part for course credit and rated all 92 “over-weighted” CIs generated by the Phase 1 participants, as will be detailed shortly.

**Stimuli for Image-Classification Task**

All stimuli for this task were the same stimuli we used in the pilot study and in Study 1.

**Procedure: Phase 1**

**Spider image presentation.** Participants were told that they were taking part in a study on multitasking. In reality, the purpose of this multitasking portion of the study was to present participants with a specific image of a spider multiple times. Over three blocks, participants saw a stream of images appear onscreen. These images included animals, people, insects, inanimate objects, etc. Each image appeared for three seconds and then automatically switched to the next image. In each block, participants were given a category (e.g., mammals; people) as well as a target object (e.g., a photograph of balloons). They were asked to simultaneously keep a running mental tally of the number of non-repeated images belonging to the category while pressing the spacebar every time they saw the target image. We included some repeated category members along with instructions to count only non-repeated category members in order to account for the fact that the spider images appeared three times per block. At the end of each block, they were asked to enter the number of category members they had counted. They were then given a new category and a new target to search for.
During each of the three blocks, a specific image of a spider appeared three separate times. This image was a full-color version of the base image we had used in the Pilot Study and in Study 1. The image was presented a total of nine times at three seconds each for a total exposure of 27 seconds.

**Image-classification task.** Once participants had completed the multitasking portion of the study, they were told that the accuracy of their memory of one of the images they had seen would be tested. The instructions were designed to appear like a form letter, where the specific object label (“tarantula”) was set apart and displayed in a different font to suggest that it had been inserted into a blank space. We hoped that this would suggest to participants that not every participant was being tested on the same image, even though all were. Participants were instructed to recall what the specific image looked like and given ten seconds to do so.

Participants were then given instructions for the image classification task. They were told that on each trial they should select which image of a tarantula (again displayed in a different font) they believed looked more like the image they saw in the first task. They were further instructed to choose the more accurate image. As with Study 1, participants did this for 450 image pairs. They were given a six-point scale to indicate whether they were very confident, confident, or would guess that the image on the left versus the right looked more like the image they had seen earlier. The question (“Which looks more like the image you saw?”) and the scale labels appeared on screen for each of the 450 trials.
Scales. After the image classification task, participants first completed the Attentional Control Scale (Derryberry & Reed, 2002). We included this scale purely because it fit with our multitasking cover story, to ensure that participants would complete the Fear of Spiders scale without guessing its relation to the image classification task. We had no predictions regarding participants’ responses to it, and it will not be discussed further. Participants then completed the Fear of Spiders scale (Szymanski & O’Donohue, 1995). The scale had good reliability in this sample ($\alpha = .96$) and all 18 items were averaged together. Once participants had completed this scale, they responded to demographic questions and were debriefed.

Assessing mental representations of the specific spider. Here, we again used reverse-correlation data-reduction techniques to generate an estimate of participants’ mental representation of the specific spider they had seen earlier in the study. As with Study 1, we generated both participant-level and group-level CIs. The 92 participant-level CIs were derived by averaging the parameters which defined each of the 450 noise patterns each participant selected. This resulted in 4092 mean parameters per participant. Noise patterns based on these mean parameters were superimposed over the base spider image to create 92 separate CIs (one per participant). We generated group-level CIs by further averaging the mean parameters for the two subgroups of participants higher versus lower in Fear of Spiders. This resulted in a pair of CIs which provided an estimate of a particular group’s mental representation of the specific spider.

As with Study 1, Phase 1 participants used a 6-point scale during the image-classification task, which allowed us to generate both dichotomous and weighted CIs for
group-level estimates, and weighted CIs for participant-level estimates. The differences between these two CI types are detailed in Study 1.

In order to boost the sensitivity of the participant-level estimates to the confidence participants had in their choices, we also generated what we are calling “over-weighted” estimates. In this case, rather than simply weighting “Very confident” trials three times, “Confident” trials twice, and “Guess” trials once, we created three separate parameter vectors per participant. The first parameter vector represented an average set of 4092 parameters based on participants’ choices on all 450 trials, regardless of confidence. The second parameter vector represented an average based on choices in the “Confident” and “Very confident” trials. The third parameter vector represented an average based on choices in the “Very confident” trials only. Once these three vectors were calculated, they were averaged together to generate each participant’s over-weighted CI. This method had the advantage of expanding the impact of the “Confident” and especially the “Very confident” trials on the ultimate CI. Note, though, that using this method means that if a participant uses a given key very infrequently (say, ten of the 450 total trials), those few selected noise patterns will together represent a third of the ultimate CI. In order to limit this possibility, we left out a given key type for a participant (“Confident” or “Very Confident”) if that participant used the key less than five percent of the time. For 29 participants, the final CI was based on two key types, and for 28 participants, the final CI was based on only one key type. For 35 participants, all three categories were involved.
Procedure: Phase 2

**Ratings of participant-level estimates.** One MTurk sample of adult volunteers rated the 92 weighted participant-level CI generated in Phase 1. A separate sample of undergraduate participants rated the 92 over-weighted participant-level CIs generated in Phase 1. Both rater samples viewed each mental representation estimate and reported how scary/threatening/frightening the tarantula looked on a 10-point scale (anchors were *not at all* and *extremely*). They then completed the Fear of Spiders scale and were debriefed.

**Ratings of group-level estimates.** Independent raters on MTurk compared the specific spider CIs based on the group of Phase 1 participants higher versus lower in fear of spiders (both dichotomous and weighted) on a series of variables. Each pair was displayed side-by-side (counterbalanced). Raters were asked to indicate in which of the two photographs the spider looked more threatening, frightening, and scarier (order randomized). They responded to the first three items on 6-point continua (anchors were *looks much more threatening/much more frightening/scarier in the photo on the LEFT (RIGHT)*). Raters also indicated in which image the spider made them feel more anxious on a 6-point continuum (anchors were *the spider in the LEFT (RIGHT) image makes me feel more anxious*). MTurk raters then indicated the extent to which they were afraid of spiders on a 4-point scale (anchors were *I am not at all afraid of spiders* and *I am extremely afraid of spiders*). They completed demographic items and were debriefed.
Results

Phase 1 participants ranged in terms of their Fear of Spiders ($M = -0.69, SD = 1.57$). The MTurk sample who rated the weighted participant-level images also varied in Fear of Spiders ($M = -0.48, SD = 1.75$), as did the undergraduate sample who rated the over-weighted participant-level images ($M = 3.52, SD = 1.50$). The sample of Phase 2 MTurk participants who rated the group-level images also ranged in their responses to the single 4-point fear of spiders item ($M = 2.44, SD = 1.02$).

Performance on the Spider Presentation Task

To assess whether participants were generally taking the presentation task seriously, we calculated the mean number of non-repeated exemplars participants counted which belonged to the category assigned to each of the three blocks. For Block 1, where the category provided was “mammals,” we presented participants with 7 mammal pictures total, 5 non-repeated. We also included 6 pictures of people, 3 non-repeated, bringing the technical non-repeated mammal total to 8. The average number of mammals participants counted was $8.17 (SD = 3.69)$. In Block 2, the category was “people” and we presented 6 people pictures total, 4 non-repeated. Participants counted an average of 4.37 people ($SD = 1.26$). In the final presentation block, the category was “insects” and we presented 4 insect pictures along with 3 presentations of the spider image (7 total insects; 4 non-repeated). Participants counted an average of 4.64 insects ($SD = 1.69$). It seemed, then, that participants followed instructions and paid enough attention during the spider presentation task to be reasonably accurate in their counts, at least on average.
Predicting the Scariness of Group-Level CIs

MTurk participants rated the group-level CIs for Phase 1 participants higher in Fear of Spiders (upper 15 percent: $n = 14$; scale value 1.39 and higher) and lower in Fear of Spiders (lower 15 percent: $n = 13$; scale value -2.39 and lower). Responses were recoded and centered around zero such that higher numbers indicated that the CI derived from participants higher in Fear of Spiders was rated as scarier, more threatening, etc. One rater was dropped from the analysis for failing a quality control item, resulting in an $N$ of 123.

**Dichotomous classification image averages.** All four ratings of the dichotomous CI pair were highly related ($\alpha = .96$), so we created a composite by averaging them together. In contrast with Study 1, in this case, Phase 2 raters reported no difference in how scary, threatening, frightening, or anxiety-inducing the higher vs. lower fear of spiders dichotomous group-level CI appeared to be, although the mean was in the predicted direction ($M = .10$, $t(122) = .71$, $p = .48$). Raters’ responses to the single fear of spiders item did not predict their ratings ($p = .39$).

**Weighted classification image averages.** When we weighted the group-level CIs by participants’ confidence in their selections, we did find the predicted effect. The four ratings of the weighted CI pair were highly related ($\alpha = .97$) and were therefore averaged together. The weighted specific spider CI based on Phase 1 participants higher in Fear of Spiders was rated as scarier, more threatening, more frightening, and more anxiety-inducing than the weighted CI based on Phase 1 participants lower in Fear of Spiders ($M = .28$, $t(122) = 2.06$, $p = .04$). Raters’ responses to the single fear of spiders item did not
predict their ratings \((p = .45)\). Again, the size of the effect for the weighted CI pair was larger than the effect for the dichotomous CI pair (dichotomous CI pair, Cohen’s \(d = .13, \ r = .06\); weighted CI pair, Cohen’s \(d = .37, \ r = .18\)). Figure 4 displays the weighted images.

![Figure 4. Weighted (bottom row) CI pair for Study 2. The CI based on participants lower in Fear of Spiders is on the left. The CI based on participants higher in Fear of Spiders is on the right.](image)

**Predicting the Scariness of Participant-Level CIs**

Since the group-level results were strongest for the weighted CIs, our sample of MTurk raters rated weighted estimates of the mental representations per Phase 1 participant of the specific spider they had seen earlier in the study. However, an HLM analysis predicting the rated scariness of these representations from Phase 1 participants’
Fear of Spiders and raters’ Fear of Spiders did not yield a significant effect of Phase 1 participant Fear of Spiders ($p = .92$).

However, ratings of the ‘over-weighted’ participant-level CIs revealed the predicted effect. We ran a two-level HLM analysis with 6808 observations based on 92 Phase 1 participants nested in 74 Phase 2 undergraduate raters. The mean rating across the 6808 observations was 5.86 and the standard deviation was 2.44. We predicted the rating of how scary/threatening/frightening a given Phase 1 participant’s mental representation of the specific spider was from Phase 1 participants’ Fear of Spiders (entered group-mean centered) at level one and from Phase 2 participants’ Fear of Spiders (entered grand-mean centered) at level 2. We modeled all coefficients as fixed if the associated error term was not significantly different from zero (using Nezlek’s (2011) recommended cutoff of $p > 0.2$).

We found the predicted main effect of Phase 1 participants’ Fear of Spiders on the rated scariness of their mental representation of the specific spider ($\gamma = .03, t(6732) = 3.59, p < .001$). The more afraid of spiders the Phase 1 participant was, the scarier and more threatening the over-weighted specific spider CI was judged to be by independent raters. We also found a main effect of the raters’ Fear of Spiders ($\gamma = .62, t(72) = 4.27, p < .001$) such that the higher in Fear of Spiders the rater was, the scarier, in general, all 92 CIs were judged to be. See Table 2 for the results of this model.
Table 2. HLM Regression Coefficients for Study 2: Predicting the Scariness of Participants' Over-Weighted Specific Spider Classification Images (CIs).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coefficient</th>
<th>SE</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (\gamma_{00})</td>
<td>5.85</td>
<td>(0.23)</td>
<td>***</td>
</tr>
<tr>
<td>Phase 2 Rater FOS (\gamma_{01})</td>
<td>0.61</td>
<td>(0.14)</td>
<td>***</td>
</tr>
<tr>
<td>Phase 1 Participant FOS (\gamma_{10})</td>
<td>0.03</td>
<td>(0.01)</td>
<td>***</td>
</tr>
</tbody>
</table>

Cross-level interactions

| Phase 1 Part. FOS X Rater FOS \(\gamma_{11}\) | -0.01 | (0.01) |

Significance: + \(p < .10\); * \(p < .05\); ** \(p < .01\); *** \(p < .001\) (two-tailed test).

Standard errors in parentheses.

Predicting the Similarity of Participant-Level CIs to Pilot ‘Scary’ Spider CI

To bolster the participant-level results stemming from raters’ judgments, we also calculated the correlation of each participant-level CI with the ‘scary’ spider estimate that was generated in the Pilot study. The higher the correlation, the more a given participant’s CI matches the CI that our pilot participants generated, on average, when asked to choose the scarier spider. To do this, we first masked out the background of the spider image. The 105419 pixels we included in each correlation comprised the body of the spider as well as the shadow underneath the spider (see Figure 5). For all three types of CI (dichotomous, weighted, and over-weighted), correlations with the ‘scary’ pilot CI
were high (dichotomous $M = .91$, $SD = .03$; weighted $M = .91$, $SD = .03$; over-weighted $M = .91$, $SD = .03$), but the extent to which each CI correlated with the ‘scary’ pilot CI varied across the participants (dichotomous range .81-.96; weighted range .81-.97; over-weighted range .81-.97). More interestingly, excluding one outlier (dichotomous Cook’s $D = .16$, next highest values = .05, .05 and .04; weighted Cook’s $D = .14$, next highest values = .06, .05, and .04; over-weighted Cook’s $D = .16$, next highest values = .06, .04, and .04), the higher in Fear of Spiders Phase 1 participants were, the more their specific spider CI correlated with the ‘scary’ spider CI from the pilot, (dichotomous $r(91) = .23$, $p = .03$; weighted $r(91) = .21$, $p = .047$; over-weighted $r(91) = .19$, $p = .068$). Even though Phase 2 raters were unable to detect meaningful differences in scariness for the dichotomous participant-level CIs, this more objective measure suggests that even the dichotomous CIs were sensitive to differences in participant Fear of Spiders.\textsuperscript{1}
Participant Accuracy

To assess whether participants were following our accuracy instructions, we used the base image, which they had seen earlier in the study, as a benchmark for accuracy. To measure accuracy in the image classification task, for each of the 450 trials we correlated each of the two image options with the base image of the spider. Recall that one of these image options had one noise pattern superimposed over it (here called the ‘original’ image) and the other option had the inverse noise pattern superimposed over it (here called the ‘inverse’ image). We created a differences score such that higher numbers indicated that the ‘original’ image correlated more highly with the base spider image than the ‘inverse’ image. This score essentially indexes the differential accuracy of the original versus the inverse image. We then created a variable which indexed the confidence with which each participant selected the ‘original’ versus the ‘inverse’ image on each of the 450 trials in the image classification task. This variable ranged on a six-point scale from -3 (indicating that participants were very confident in selecting the ‘inverse’ spider image on a given trial) to +3 (indicating that participants were very confident in selecting the ‘original’ spider image on a given trial). For each participant, we then calculated the correlation across trials between their response and the differential accuracy of the original versus inverse image.

Our first question was whether, overall, participants seemed to be making decisions driven by accuracy. If so, then the correlation between each participant’s response in selecting the ‘original’ versus the ‘inverse’ image and the differential accuracy of the ‘original’ versus the ‘inverse’ image should be positive. For 65.2% of
participants, this correlation was positive, and in addition, when tested against zero, the average correlation was significantly positive (\(M = .03, \text{SD} = .08, t(91) = 2.92, p = .004\)). Participants, overall, seemed to select the more accurate image.

Did this response-accuracy metric vary by participants’ fear of spiders? We did not find evidence of this. The extent to which participants confidently selected the more accurate image in each trial (specifically, the per-participant correlation between confidence in selecting the ‘original’ image and the relative accuracy of the ‘original’ image) did not relate to participants’ spider fear (\(r(92) = .10, p = .34\)). This suggests that participants higher in fear of spiders were no more or less accurate in their selections during the image classification task than participants lower in fear of spiders.

Another way to examine whether participant accuracy varied as a function of their fear of spiders is to assess accuracy at the level of each participant’s CI. To get an idea of how accurate participants’ mental representation of the specific spider actually was, we correlated each participant’s CI (dichotomous, weighted, and over-weighted) with the original photograph of a spider we used as a base image. Given that the base image forms part of the CI, these correlations were very high (dichotomous, \(M = .97, \text{SD} = .004\); weighted, \(M = .97, \text{SD} = .004\); over-weighted, \(M = .97, \text{SD} = .004\)). Nevertheless, the variability across participants was meaningful, as evidenced by a significant correlation between this CI-accuracy metric and the response-accuracy metric reported in the previous paragraphs (dichotomous, \(r(92) = .23, p = .025\); weighted, \(r(92) = .30, p = .003\); over-weighted, \(r(92) = .33, p = .001\)). Importantly, though, like the response-accuracy metric, the CI-accuracy metric did not correlate significantly with participants’ fear of
spiders (dichotomous, $r(92) = .005, p = .96$; weighted, $r(92) = -.07, p = .50$; over-weighted, $r(92) = -.12, p = .28$).

**Discussion of Study 2**

The findings from Study 2 provide further evidence that even when asked to recall a specific spider image from an earlier task, participants’ mental representations of that spider were influenced by their relative fear of spiders. Participants with higher fear seemed to mentally represent that specific spider in a way that was judged by objective observers as looking scarier and more threatening than the representation of that same specific spider in the minds of participants lower in fear of spiders. We found evidence of this both at the group level and at the participant level. At the group level, we found that an average CI from participants higher in fear was rated as scarier and more threatening-looking than an average CI from participants lower in fear of spiders. At the participant level, we found both subjective and objective evidence that fear of spiders related to the scariness of an individual’s mental representation of a specific spider. Using subjective ratings, we found that with over-weighted CIs, the higher in fear of spiders, the more scary and threatening a participant’s CI appeared to be. Adopting a more objective approach, we also found evidence that participants’ CIs were more strongly correlated with the ‘scary’ CI generated in the pilot study if those participants were higher in fear of spiders. In sum, participants’ construal of a specific object was biased in line with the extent of their fear.
Accuracy

Participants appeared to be accurate in terms of their choices during the image classification task. Specifically, the more similar a given image was to the base spider they had seen earlier, the more confidently they chose that image. The task was, in other words, a doable one. Participants were not responding completely randomly and they appeared to take the task seriously. These data suggest that participants did not redefine the instructions during the task, selecting instead, for example, the scarier of the two spiders or the more typical-looking of the two. Rather, they tended to select the more accurate spider – the spider that looked more like the specific spider they had seen earlier in the study.

In the introduction, we discussed research which suggests that spiders draw the attention of people high in spider fear and hold it for longer periods than people lower in spider fear (e.g., Van Strien et al., 2009). Might this explain our effects in Study 2? Are participants higher in FOS simply paying more attention to the spiders in the first place, and holding a mental representation that is more accurate than the mental representation for participants lower in FOS? Perhaps the difference in scariness we are finding is due to high-fear participants being more accurate and low-fear participants misremembering the spider as looking less scary than it actually was. First, given our paradigm, there is no reason to believe high FOS participants focused on the spiders in the presentation task any longer than low FOS participants. In an eye-tracking study, Rinck and Becker (2006) found that while spider fearfuls did spend more time looking at a picture of a spider during the first 500 milliseconds of presentation, this time became equivalent between
high and low-fear groups for the next second, after which the pattern reversed. In other words, participants higher in FOS first directed their attention to the spider picture, but then actively avoided the picture. Our stimuli were presented for three seconds each. Given the timing of attendance-avoidance reported in Rinck et al. (2006), it is possible that our participants higher in FOS did not differ in the mean duration of exposure to the spider images.

Although we don’t have direct evidence of this, we do have data that support this possibility. Recall that both at the level of trial selection in the image classification task and at the level of each participant’s CI, fear of spiders did not relate to accuracy. If participants higher in fear of spiders looked for a longer amount of time at the specific spider, they should have been more accurate in their per-trial selections and their CI should have correlated more highly with the base image. That is not what we found. While these were not direct measures of the attention these participants gave to the spider image (or of their general exposure), these results do not fit with the hypothesis that participants higher in fear of spiders attended more to the spider stimuli.

It seems to be the case, then, that while participant accuracy did not covary with participant fear, participants were, in general, accurate. How does this fit with our findings that participants higher in fear produced a CI which looked scarier than the CI from participants lower in fear? If participants higher in fear of spiders were the only ones construing the specific spider in a biased fashion, then fear should have correlated negatively with accuracy. If participants lower in fear of spiders were the only ones construing the specific spider in a biased fashion, then fear should have correlated...
positively with accuracy. Neither was the case, though, suggesting that perhaps both types of participants mentally represented the spider in a biased way, higher-fear participants as looking more scary than it actually was, and lower-fear participants as looking less scary than it actually was.

It is unclear at what point this bias comes into play. It could be that participants higher in fear construe the spider image as looking scarier while they are encoding it. It could also be that these participants recall the spider image in a biased fashion. Both mechanisms could be at play. This study speaks to the fact that whatever the mechanism, the way in which participants represent the specific spider is biased in line with their attitudes.
Chapter 5: General Discussion

Overall, we accomplished our three main goals. First, we successfully applied RCIC techniques to non-facial, complex stimuli (spiders). By modifying the RCIC task, presenting participants with single images (Dotsch et al., 2008) and a confidence scale (Mangini & Biederman, 2004), we were able to weight the CIs we generated by participant confidence, allowing us to get a clearer estimate of the actual representation in the minds of participants. Second, we were able to employ the modified RCIC technique to index meaningful variation in how scary/threatening/frightening the mental representation of these spiders appeared to be, as rated by objective judges. Specifically, participants higher in fear of spiders generated mental representations which were rated as scarier than the mental representations of participants lower in fear of spiders. Third, we found evidence that the variation in the scariness of these mental representations occurred for both a prototypical spider (Study 1) and for a specific spider participants saw earlier in the study (Study 2). This suggests that fear of spiders has power to bias the way in which participants mentally represent, or construe, an object relevant to that fear, both as a general category and when recalling a specific exemplar.
Why Weight?

Given that other RCIC research has found effects when participants’ classifications are dichotomous, why did we decide to incorporate a confidence scale to weight the ultimate CIs? Mangini & Biederman (2004) in their single-image forced-choice RCIC task incorporated a confidence scale, and we thought it would be useful to adapt this idea to the two-images forced-choice paradigm. A confidence scale was especially useful for these particular studies because participants are not likely to be as good at discriminating between and classifying spiders as they are at discriminating between and classifying faces. We therefore wanted to leverage trials on which participants reported more confidence in their classifications when building the CIs rather than weight all trials, even trials on which participants saw no difference at all between the stimuli and were truly guessing, equally. Based on our data, it seems we were justified in weighting our group-level images, especially in Study 2, which presented participants with a particularly difficult task of recalling a specific spider image. In this study, we found the predicted effect (that is, that the group-level CI based on participants higher in fear of spiders was rated as scarier, more threatening, etc. than the group-level CI based on participants lower in fear of spiders) only for the weighted image pair.

This rationale also applies to our use of “over-weighted” subject-level images in Study 2. Recall that while we found group-level effects in Study 2 with weighted images, we only found subject-level correlations when we weighted subject-level CIs even more heavily by the more confident trials. We think this is justified considering not only the difficulty of the RCIC task for participants in Study 2 but also the fact that finding
subject-level correlations in the first place is difficult. Imhoff et al. 2013 found weaker subject-level effects compared to group-level effects and suggested that this might be because subject-level images are noisier and less clear than group-level images. This is something we have observed in our own CIs. One way to improve the clarity of CIs is to boost the number of trials used to generate them – this likely a reason group-level CIs, which are based on many trials from multiple people, seem to show more consistent differences. Indeed, Dotsch et al. (2008) did not find subject-level effects of implicit prejudice on face criminality and trustworthiness ratings until they increased the number of trials in a replication from 390 to 770, which served to increase the CI quality. It is possible that increasing the number of trials to begin with, especially with such a difficult task and such a subtle effect, would have increased the likelihood of finding our predicted effects, but weighting and over-weighting served the same purpose while keeping demand on participants relatively low. Weighting the CIs increased the number of trials included in each participant’s CI, making each CI clearer and less noisy, and making any subject-level effects easier to find.

Over-weighting in Study 2 also seemed justified based on previous single-image forced choice work in which confidence scales were used (Mangini & Biederman, 2004; Kontsevich & Tyler, 2004). These studies limited their CIs to trials on which participants provided extreme responses only, dropping other trials completely. Mangini & Biederman (2004), for instance, asked participants to respond “Probably Travolta,” “Possibly Travolta,” “Possibly Cruise” or “Probably Cruise,” but based their CIs on “Probably” trials only. Kontsevich & Tyler (2004) based their CIs for Mona Lisa on the
extreme classifications (“Sad” and “Happy”), leaving out trials based on moderate classifications (“Slightly Sad” and “Slightly Happy”). Neither paper reports results including all possible trials, and it is unclear whether or not using the extreme trials was necessary to achieve their effects. Since in many cases, our participants guessed on a majority of the trials, we did not want to drop those trials completely. We felt weighting the trials was a good compromise. However, given the difficulty of recalling a specific spider in Study 2, over-weighting shifts our CIs to resemble even more those CIs used by these earlier researchers by boosting the impact of the confident trials while still maintaining the signal provided by the guess trials.

**Visual Representations?**

Note that we have been careful to use the term “mental representation” when discussing the output from the RCIC tasks. It is still debatable whether RCIC measures an actual visual representation in memory or is simply a construction based on semantic information. We would argue that both are probably true to some extent, and that it likely depends on the object participants are being asked to construct during the task. For instance, it seems more intuitively plausible to hold a visual image in memory of a specific spider one has seen before than it is to hold a visual image of a prototypical tarantula. Perhaps in Study 1, then, the RCIC task might be getting more at a construction based on participant fear, whereas Study 2 might be getting at more of a visual representation of the presented spider, albeit biased by participants’ fear. Some researchers have argued that it may be possible to have a visual representation of even a
prototype, though, and that RCIC methods could be estimating such a visual representation (Dotsch et al., 2013).

Mangini & Biederman (2004) caution against running too far with the argument that CIs reveal visual representations. The form the CIs take in any RCIC study does not vary freely to fit whatever form is in the participant’s head. It is determined at least in part by the function used to generate the noise (sinusoid), the regression function for the 4092 parameters (linear) and the base image (the particular photograph of a spider). However, this does not necessarily mean RCIC is incapable of estimating a participant’s visual representation. These researchers further suggest that sinusoidal noise is not an arbitrary choice; it is matched to the way neurons in early cortical stages represent visual stimuli. In addition, while presenting participants with noise alone rather than noise superimposed over a given base image might be better in terms of allowing the representation to come about on its own, the number of trials to do this, especially with such a complex referent as a spider, would be prohibitive. Having a base image to constrain variability defines the category (spider) but allows what makes a spider particularly scary to vary. And to reiterate a point made in the introduction, since the base image is constant across trials, it does not predict variation in participant responses. The noise does.

Whether people are even capable of holding a visual image in memory has been a long-standing debate in the field for decades (e.g., Pylyshyn, 1973; Kosslyn & Pomerantz, 1977; see Kosslyn, 2005 for a brief review). Note, though, that even proponents of visual imagery do not hold that a visual image is a photograph in the head;
rather, it is, as Kosslyn & Pomerantz (1977) put it, a “quasi-pictorial, spatial [entity] resembling those evoked during perception itself” (p. 57). There is evidence, though, that visual imagery and perception activate similar cortical regions (e.g., O’Craven & Kanwisher, 2000; see Kosslyn, Ganis & Thompson (2001) as well as Kosslyn, Thompson & Ganis (2006) for reviews). It is not our goal to prove that visual representations exist. However, it seems reasonable that if, in fact, one can store a visual representation in memory, which is analogous to the experience of perceiving a given object, the RCIC method provides a means of estimating that visual representation.

**Future Directions**

In future, it would be interesting to manipulate participants’ fear of spiders directly and see if that impacts their mental representations. This is certainly possible considering work by Rachman & Whittal (1989) demonstrating that post-treatment, patients with spider phobia subjectively rate spiders as less hairy, ugly, and threatening compared to pre-treatment ratings. A clinical intervention might similarly serve to reduce the scariness of patients’ mental representations of spiders. If reducing spider phobia also changes the scariness of the mental representation of a prototypical spider, such a change might make it even easier for patients to consider approaching spiders in the future. A change in the mental representation might mediate the effects of the clinical intervention on behavior. Further, clinical interventions might be able to use RCIC techniques to assess the effectiveness of a given intervention. If fear of spiders has been effectively reduced, patients should mentally represent spiders in less threatening, less scary-looking ways.
Conclusion

The present findings demonstrate that RCIC methods can be adapted to complex, non-facial stimuli. Not only that, but they provide further evidence that the state of an observer, specifically, that observer’s trait level of fear, can bias the way in which he or she mentally represents a fear-related object (a spider). Prior literature has necessarily relied in participant’s subjective reports of their perception of spiders; that is, the size of spiders, their speed, closeness, etc. RCIC methods provide a more direct estimate of the exact mental representation of a given spider (either prototypical or specific).
References


Kosslyn, S. M., Thompson, W. L. & Ganis, G. (2006). *The case for mental imagery*. Oxford University Press.


Footnotes

¹Note that we checked this correlation for Study 1 and found a null relationship between Fear of Spiders and the correlation of participants’ CI with the ‘scary’ spider CI from the pilot. While we initially anticipated finding such a relationship, we now think it is reasonable that we did not find it in the case where participants were asked to select which spider looked more like a typical tarantula. There is likely much more variation in the CIs in Study 1 compared to the CIs in Study 2, where participants are asked which image looks more like the specific photograph (which also served as the base image for all stimuli). When imagining a typical spider in Study 1, individuals higher in fear of spiders do seem to imagine a scarier one, but it may be the case that the specific visual features that make that prototypical tarantula scary may vary considerably from one participant to another. This CI variation might disguise the pixel-correlation relationship with fear of spiders. Subjective ratings of scariness are not as dependent in Study 1 on the participant’s CI looking exactly like the “scary” spider from the pilot study – a spider can look scary for a number of different reasons – so it makes sense that we find subjective effects in both studies but objective effects only in Study 2.