Thinking & Reasoning, 2014
Vol. 20, No. 3, 315–332, http://dx.doi.org/10.1080/13546783.2013.811442

Integrating information from multiple sources: A metacognitive account of self-generated and externally provided anchors

Keith W. Dowd, John V. Petrocelli, and Myles T. Wood

Department of Psychology, Wake Forest University, Winston-Salem, NC, USA

Estimates of unknown quantities are influenced by both self-generated anchors (SGAs) and externally provided anchors (EPAs; e.g., the advice of others). It was hypothesised that people use the degree of similarity between these anchors to render final responses. Thus we tested predictions drawn from metacognitive accounts of anchoring using procedures similar to the traditional anchoring paradigm. In a single experiment we manipulated SGA–EPA similarity, EPA level, and EPA source credibility. Results showed that the relationship between SGA–EPA similarity and the decision weight given to the EPA, relative to the weight given to the SGA, depended on source credibility. Bolstering the metacognitive framework, participants were most confident about their final responses when their SGA was similar to the EPA and the EPA came from a highly credible source. These results support a metacognitive account of the anchoring heuristic.

Keywords: Anchoring heuristic; Judgement and decision making; Metacognition.

In everyday life people frequently estimate unknown numerical quantities. For example, one may estimate the value of real-estate or guess how much sugar to mix into a recipe. One strategy for estimating such quantities is to use estimates made available by the context (i.e., anchors). Although people are sensitive to estimates that are too high or too low, research has shown they tend to generate values that are too close to these estimates (Chapman & Johnson, 2002; Tversky & Kahneman, 1974; Yaniv & Milyavsky, 2007).

The authors are most grateful to Catherine E. Seta and John J. Seta for their feedback on an earlier version of this work as well as their assistance with important calculations.

Address correspondence to: John V. Petrocelli, Department of Psychology, Wake Forest University, P.O. Box 7778, Winston-Salem, NC 27109, USA. E-mail: petrocjv@wfu.edu

That is, when estimating unknown numerical quantities people are influenced by the presence of unrelated numbers, or "anchors".

In the current research, we define an *externally provided anchor (EPA)* as an anchor provided by an external source within the environment. Although research has demonstrated that EPAs have a significant influence on human judgement and decision making (Chapman & Johnson, 1999; Jacowitz & Kahneman, 1995; Ritov, 1996), few process accounts have examined the role of the *self-generated anchor (SGA)*. SGAs are internally generated approximations of an estimate.

One exception to the neglect of SGAs is the work of Epley and Gilovich (2001, 2004, 2005, 2006). In fact, they argued that SGAs, rather than EPAs, demonstrate actual adjustment. Our view of the anchoring heuristic is that an EPA typically influences judgement through a relatively conscious and deliberate process. However, we contend that the generation of SGAs (or an implicit, plausible range) often comes to mind rather automatically. For example, most people are aware that George Washington was elected President of the United States after the Declaration of Independence was written (i.e., 1776), but are less clear about the actual year he became U.S. President. Further, as the work of Epley and Gilovich has shown, people appear to be aware that their SGAs are not the correct answer from the beginning of the process, adjusting from these values until more plausible final responses are reached.

However, much of the research using SGAs has employed "common knowledge" questions and assumed that all participants activate the same SGA (e.g., Epley & Gilovich, 2001). The current research examines the weight that people give to both an EPA and their personal SGA in their final judgements.

McElroy and Dowd (2007) suggested that some individuals (e.g., individuals low in the openness personality dimension) could be less susceptible to EPAs and rely more on internally generated information to respond to traditional anchoring questions. Here we suggest that people use EPAs to aid in confirming or disconfirming their SGAs and use contextual information to integrate both the SGA and EPA to form a final estimate. Thus we tested a metacognitive account of the anchoring heuristic that involves three stages.

First, we assume that when people are asked to estimate uncertain quantities they automatically generate an SGA (or at least an implicit range of plausible values) that becomes part of the context for their subsequent judgements. Support for this notion is found from research that indicates a strong tendency for people to *correctly* make upward adjustments from low anchors and downward adjustments from high anchors (Chapman & Johnson, 2002; Epley & Gilovich, 2001, 2004, 2005, 2006; Tversky & Kahneman, 1974). In fact, people are generally accurate in knowing whether an anchor is too high or too low for it to be plausible, implying that an implicit, plausible range of values is activated in such cases. For example, consider the number of books that have been written about George Washington; people seem to know that the correct answer is not "zero" or "one million". It seems reasonable to expect such implicit ranges to operate as standards of comparison for judgements and for these implicit ranges to be highly accessible when first encountering traditional anchoring questions.

Second, we build on Yaniv's (2004a, 2004b) advice and consensus research to propose that SGAs serve as important reference points for judging the feasibility and accuracy of EPAs. We propose that people are generally more confident in their final responses when the similarity between their SGAs and EPAs is high than when it is low. Essentially, SGA–EPA similarity may signal important consensus information, and greater consensus information can increase certainty in judgements (Petrocelli, Tormala, & Rucker, 2007). We also expect SGA–EPA similarity to play a role in anchoring effects because it is likely to be positively correlated with the perceived plausibility, and negatively correlated with the perceived extremity, of an EPA. In fact, studies that have manipulated plausibility (Wegener, Petty, Detweiler-Bedell, & Jarvis, 2001) and extremity (Mussweiler & Strack, 2001) have found these variables to be important to anchoring.

Third, aspects of the EPA, such as source credibility, have been shown to moderate anchoring effects. EPAs are more influential when they come from high rather than low credible sources (see Wegener, Petty, Blankenship, & Detweiler-Bedell, 2010; Wegener et al., 2001). Consistent with information integration research and theory (Anderson, 1970, 1981, 1982), we propose that people use their SGA, in addition to other relevant information (e.g., EPA source credibility), to render a final estimate. These possibilities are directly examined in this research.

Specifically, we hypothesised that when SGA–EPA similarity is high, people may infer from it that their SGA and the EPA are reasonable. In such cases we expect the decision weight placed on the EPA to be relatively large. However, when SGA–EPA similarity is low, people may infer from it that their SGA (or the EPA) is less reasonable. Because we believe it is unlikely that people will abandon their SGA in such cases, we expect the decision weight placed on the EPA to be relatively small when SGA–EPA similarity is low than when it is high. If this reasoning is correct, it should be possible to demonstrate that SGA–EPA similarity affects the decision weight placed on the EPA using procedures similar to the traditional anchoring paradigm.

EXPERIMENTS 1A AND 1B

One concern with the proposed metacognitive model is the question of whether or not SGAs are spontaneously generated or used only when they are made salient by an experimental paradigm. An additional possibility is that simply requesting people to explicitly report their SGAs might have an effect on final responses. Before testing our primary hypotheses, we first sought to determine whether or not SGAs are spontaneously generated and whether or not explicitly stating one's SGA has an effect on one's final response. In Experiment 1A we compared the SGAs of participants who reported their SGAs before being presented with an EPA with those of participants who were asked to recall their SGAs after being presented with an EPA and reporting their final responses. In Experiment 1B we compared the final responses of participants who were asked to report their SGAs to those of participants who were not asked to report their SGAs.

We reasoned that measuring SGAs for the purposes of making valid conclusions from an experiment was possible to the extent that the final response is not affected by simply reporting one's SGA and to the extent that the SGA is not affected by the period in time in which it is requested (either before or after the presentation of the EPA). Furthermore, we reasoned that the generation of SGAs is spontaneous to the extent that we demonstrate that final responses are not affected by the expression of an SGA.

Method

Participants and design. All participants ($N_{1A} = 56$, $N_{1B} = 54$) were enrolled in introductory psychology courses, were recruited through an electronic participant pool, and received credit.

Experiment 1A employed a two-factor design whereby we manipulated the EPA (high vs low) as well as the time in which the presentation of the EPA was presented (either before or after the expression of the SGA); we examined both SGA and final response as dependent variables. Experiment 1B also employed a two-factor design in which we manipulated the EPA (high vs low) as well as whether or not participants were asked to express their SGA; we examined final response as the dependent variable.

Procedure. In both Experiments 1A and 1B participants were provided with a brief oral introduction to the experiment and then escorted to a cubicle with a personal computer. All of the instructions and stimuli were presented using MediaLab v2006 Research Software (Jarvis, 2006). All participants in both experiments were first asked to read the following statement: "Consider the population of Chicago, IL (in millions)."

In Experiment 1A participants were randomly assigned to one of two EPA conditions (high vs low) and one of two SGA report conditions (before vs after EPA). In the *report SGA before EPA* condition after participants read the statement they were then asked to answer the SGA question: "What is your estimate of the population of Chicago, IL (in millions)?" Next they were asked to respond to the comparative assessment/EPA question: "Is the population of Chicago, IL more or less than 5 [.2] million?" Finally participants reported their final estimate of the population of Chicago, IL (in millions). In the *report SGA after EPA* condition after participants read the statement they were immediately presented with the comparative assessment/EPA question and were asked to provide their final estimate. They were then asked to recall from memory their SGA.

In Experiment 1B participants were randomly assigned to one of two EPA conditions (high vs low), as well as to one of two SGA report conditions (no report vs report). Participants assigned to the *no report of SGA* condition followed the same procedure as that of the report SGA after EPA condition of Experiment 1A, but they were not asked to recall their SGA. In the *report SGA* condition participants were asked the same series of questions as participants assigned to the report SGA before EPA condition of Experiment 1A.

Results and discussion

In Experiment 1A we analysed mean differences using standard *t*-tests, and also used Bayesian *t*-tests to examine evidence for the null hypotheses that both SGAs and final responses did not differ by SGA report condition (Rouder, Speckman, Sun, Morey, & Iverson, 2009). This method transforms a *t*-value into a Bayes factor, which is the ratio of the probabilities of the data given the null and alternative hypotheses (see Matthews, 2011). The Bayes factor can be used to revise beliefs about the probabilities of the null and alternative hypotheses. Following Rouder et al. (2009), the alternative hypothesis is represented by a distribution of plausible effect sizes.

The SGA for participants who reported their SGA before being presented with the EPA averaged 5.45 ($SD = 5.06$), and it did not differ from that of their counterparts who reported their SGA after being presented with the EPA ($M = 5.32$, $SD = 5.33$), $t(54) = .10$, $p = .92$. We then converted this result to a scaled JZS Bayes factor, $B_{01} = 4.97$; this factor indicates that the data are almost five times as probable under the null hypothesis (no difference) as under the alternative hypothesis.

The final response for participants who reported their SGA before being presented with the EPA averaged 3.93 ($SD = 4.07$), and it did not differ from that of their counterparts who reported their SGA after being presented with the EPA ($M = 4.82$, $SD = 4.82$), $t(54) = -.75$, $p = .46$. We converted this result to a scaled JZS Bayes factor, $B_{01} = 3.88$; thus this factor indicates that the data are almost four times as probable under the null hypothesis as under the alternative hypothesis. According to Raftery (1995), the results for the SGA and final response data provide "positive" evidence for the null hypothesis.

Similar to our analytical approach for Experiment 1A, we examined evidence for the null hypothesis that final responses did not differ by reporting or not reporting one's SGA. The final response for participants who reported their SGA averaged 3.13 ($SD = 2.60$), and it did not differ from that of their counterparts who did not report their SGA ($M = 3.08$, $SD = 2.18$), $t(52) = .08$, $p = .94$. We again converted this result to a scaled JZS Bayes factor, $B_{01} = 4.90$; thus the data are almost five times as probable under the null hypothesis as under the alternative hypothesis. According to Raftery (1995), this result also provides "positive" evidence for the null hypothesis.

Thus the results of our pilot tests suggest that SGAs can be measured in an experimental setting without influencing the SGA itself or the final response. These results also suggest that the generation of SGAs is a spontaneous mental activity.

EXPERIMENT 2

If people use a metacognitive process involving SGA–EPA similarity, it should also be possible to demonstrate that specific information about the EPA influences final responses. In the current experiment we manipulated the source credibility of the EPA and hypothesised that people are most influenced by EPAs that come from highly credible sources. However, it is important to note that we expect SGA–EPA similarity to be the primary mechanism that affects the distance of final responses from the EPA because SGA–EPA similarity should be relatively more proximal and salient than peripheral features, such as the source's credibility. Thus EPA source credibility should play a bigger role in final responses to anchoring items when SGA–EPA similarity is low than when it is high. To test this prediction we calculated the decision weight given to the EPA as our dependent variable for both high and low EPA items. Consistent with our reasoning above, we expected to find evidence for a two-way interaction between SGA–EPA similarity and EPA source credibility.

Finally we also made predictions about the confidence that people have in their final responses. Jacowitz and Kahneman (1995) showed that people treat EPAs as useful information, as evidenced by their anchored participants tendency to report greater confidence in their final estimates than unanchored participants. We reasoned that people treat features of the EPA (e.g., its source and similarity to the SGA) as useful information as well. We hypothesised that when SGA–EPA similarity and EPA source credibility were both high, people would feel justified in holding relatively more confidence in their final responses than in any other condition. Such findings would further support our position that final responses to anchoring questions result from metacognitions involving thoughts about one's SGA and the EPA.

Method

Participants and design. A total of 214 participants enrolled in introductory psychology courses participated in the current experiment. All participants were recruited through an electronic participant pool and received credit towards their research course requirements in exchange for participation.

The current experiment employed a 2 (Source Credibility: high vs low) × 2 (SGA–EPA Similarity: high vs low) × 2 (EPA Item: high vs low) × 2 (EPA Item Order: high first vs low first) mixed factorial design, with EPA Item serving as the only within-participants variable. The primary dependent variable was the calculated decision weight given to the EPA (relative to the SGA).[1] We also examined self-reported confidence in participants' final response as a dependent variable.

Procedure. Participants were provided with a brief oral introduction to the experiment and then escorted to a cubicle with a personal computer. All of the instructions and stimuli were presented using MediaLab v2006 Research Software (Jarvis, 2006).

Participants were randomly assigned to one of eight experimental conditions. These conditions varied from one another only with respect to the three between-participants variables (the third between-participants variable counterbalanced the order of the high vs low EPA).

Participants were led to believe that a recent survey was administered to several hundred senior college students. They were then presented with the first of two anchoring items: "If a plane leaves from Greensboro, NC and flies directly to New York, NY, then flies directly from New York, NY to Detroit, MI, and then flies directly from Detroit, MI to Denver, CO, what is the total distance (in miles) that the plane has travelled?" The second item was of the same structure but included the following cities: Dallas, TX; Buffalo, NY; Indianapolis, IN; and Atlanta, GA. It is important to note

[1] We considered a number of dependent variables. One might expect the most straightforward dependent variable to simply be the difference between the EPA and the final estimate. However, because the EPA was calculated on the basis of the SGA artificial effects might be detected. For instance, even if SGAs and final estimates did not differ across all four cells of the EPA × Similarity portion of the design (e.g., all fours cells = 3000) both before and after the EPA is presented, we would still find an EPA Item × SGA–EPA Similarity interaction. Using the final estimate as the dependent variable was also considered. However, we feel that this measure largely ignores an important aspect of our approach (i.e., the SGA) and the fact that the EPA was calculated on the basis of the SGA.

that the actual total distances for the two anchoring items are equal (i.e., 2067 miles). Before responding to the second anchoring item, participants completed an unrelated set of self-report questionnaires (taking less than 10 minutes to complete) in order to erase their short-term memory for the first anchoring item task. After completing the second item, participants were debriefed.

Self-generated anchor. For each anchoring item participants were given 45 seconds to consider the item before being required to report an absolute estimate (between 500 and 9999 miles) about the total distance travelled. This estimate served as their SGA.

SGA–EPA similarity and calculated EPA. SGA–EPA Similarity was manipulated in the high EPA condition by adding 7% (high similarity condition) or 35% (low similarity condition) to the SGA reported by the participant. In the low EPA condition SGA–EPA Similarity was manipulated by subtracting 7% or 35% from the SGA. These values were calculated by the software and subsequently presented as the EPAs.

Source credibility. Coupled with each EPA was information about its source. Specifically, participants were informed that a randomly selected college senior, studying either fine arts (low credibility) or geography (high credibility), had responded to the anchoring item. The source's response was presented to participants as the EPA (calculated using the SGAs provided by participants).

Comparative assessment and final responses. For each anchoring item participants were asked to indicate whether they believed that the actual distance travelled was more or less than the EPA before being required to render their final response.

Confidence in final responses. Finally, for each anchoring item participants were asked to rate the degree of confidence they had in their final response using a 7-point scale, with *not at all confident* (1) and *extremely confident* (7) as the anchor labels.

Results and discussion

Consistent with expectations, a 2 (EPA Item: high vs low) × 2 (EPA Item Order: high first vs low first) repeated-measures analysis of variance (ANOVA) revealed a main effect of EPA Item (high vs low) on final estimates, such that participants recorded greater estimates for the high EPA item ($M = 3137$, $SD = 2147$) than they did for the low EPA item ($M = 2325$, *SD* = 1390), $F(1, 212) = 37.74$, $p < .001$. EPA Item Order did not affect final responses, $F(1, 212) = .01$, *ns*; EPA Item × 2 EPA Item interaction: $F(1, 212) = .88$, *ns*. Thus EPA Item Order was excluded from all subsequent analyses.

Decision weight given to the EPA in final response. Consistent with information integration research and theory (Anderson, 1970, 1981, 1982), as well as that of the advice-giving literature (Yaniv, 1997, 2000b; Yaniv & Foster, 1997), we used a weighted averaging approach to calculate our primary dependent variable. Because we were interested in determining the extent to which a participant's final estimate might be influenced by both anchors (i.e., SGA, EPA), we first set the final estimate equivalent to the following ($[W_{SGA} \times SGA] + [W_{EPA} \times EPA]) / (W_{SGA} + W_{EPA})$; $W_{subscripts}$ represent the decision weights given to the two different anchors and essentially indicate the degree to which the final estimate is a function of the two anchors. Ensuring that $W_{SGA} + W_{EPA} = 1$, we solved for W_{EPA}, which was equivalent to: (Final Estimate – SGA) / (EPA – SGA); making the W_{SGA} equivalent to $1 - W_{EPA}$. Easily interpreted, we focused our analyses on the W_{EPA}. Specifically, a positive W_{EPA} value indicated that the participant adjusted his/her final estimate closer to the EPA. As the W_{EPA} exceeded a value greater than that of .50, a participant's final estimate was closer to the EPA than his/her SGA; as the W_{EPA} came closer to zero, a participant's final estimate was closer to his/her SGA than the EPA. A negative W_{EPA} value indicated that not only did the EPA fail to receive any weight, as captured by the final estimate, but the final estimate was polarised in the direction of the SGA (and was even further away from the EPA). These data were then analysed using a 2 (Source Credibility: high vs low) × 2 (SGA–EPA Similarity: high vs low) × 2 (EPA Item: high vs low) mixed analysis of variance.

From this analysis we observed a main effect of EPA Item, $F(1, 210) = 4.02$, $p < .05$, such that participants gave greater weight to the EPA in response to the high EPA ($M = .24$, $SD = .97$) than the low EPA ($M = .09$, $SD = .97$). However, this effect was qualified by the predicted SGA–EPA Similarity × Source Credibility interaction, $F(1, 210) = 41.74$, $p < .001$. These effects were not significantly qualified by the three-way interaction, $F(1, 210) = 3.01$, $p = .08$ (see Figure 1). No other statistically significant effects emerged from our analysis. Because the three-way interaction was marginally significant we chose to analyse the high and low EPA items' data separately using the error term and degrees of freedom from the three-way ANOVA.

High EPA item data. For the decision weight given to the high EPA item data, a significant main effect of SGA–EPA Similarity emerged, $F(1, 210) = 16.12$, $p < .001$; on average, participants placed less weight on

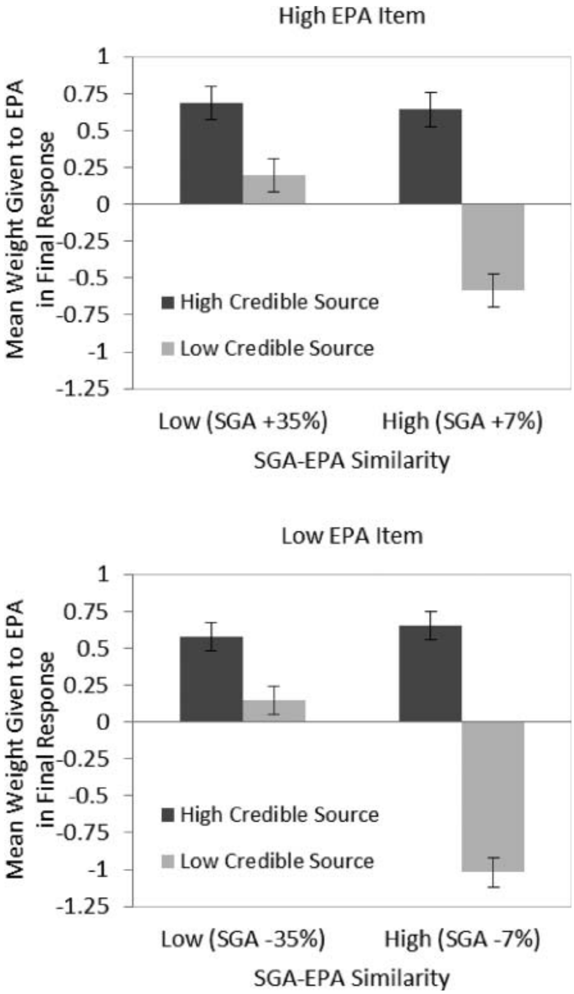

Figure 1. Mean weights given to the EPA in the final response by EPA item, SGA–EPA similarity, and source credibility.

the EPA when the SGA and EPA were similar (M = .03, SD = 1.32) than when they were not (M = .44, SD = .29). A significant main effect of Source Credibility also emerged as expected, $F(1, 210) = 69.18$, $p < .001$; on average, participants placed more weight on the EPA when the credibility of the source of the EPA was high (M = .67, SD = .36) than when it was low (M = –.19, SD = 1.18). However, these effects were qualified by the predicted SGA–EPA Similarity × Source Credibility interaction, $F(1, 210) = 12.73$,

$p < .001$. When SGA–EPA Similarity was low, participants placed significantly greater weight on the EPA when Source Credibility was high than when it was low, $t(210) = 3.37$, $p < .001$. This same pattern was found when SGA–EPA Similarity was high, $t(210) = 8.37$, $p < .001$. Also, when the EPA came from a highly credible source, no difference was found in the weight placed on the EPA when comparing high and low SGA–EPA Similarity, $t(210) = .31$, *ns*. However, when the EPA came from a source low in credibility, participants placed significantly less weight on the EPA when SGA–EPA Similarity was high than when it was low, $t(210) = 5.38$, $p < .001$.

Low EPA item data. A similar pattern of results was found for the decision weight given to the low EPA item data. As expected, a significant main effect of SGA–EPA Similarity emerged, $F(1, 210) = 28.02$, $p < .001$; on average, participants placed less weight on the EPA when the SGA and EPA were similar ($M = –.18$, $SD = 1.29$) than when they were not ($M = .36$, $SD = .27$). A significant main effect of Source Credibility also emerged, $F(1, 210) = 103.95$, $p < .001$; on average, participants placed more weight on the EPA when the credibility of the source of the EPA was high ($M = .62$, $SD = .45$) than when it was low ($M = –.42$, $SD = 1.06$).

However, again these effects were qualified by the predicted SGA–EPA Similarity $\times$ Source Credibility interaction, $F(1, 210) = 36.27$, $p < .001$. When SGA–EPA Similarity was low, participants placed significantly greater weight on the EPA when Source Credibility was high than when it was low, $t(210) = 2.96$, $p < .01$. This same pattern was found when SGA–EPA Similarity was high, $t(210) = 11.41$, $p < .001$. Once again we found that when the EPA came from a highly credible source, no difference in weight placed on the EPA was found when comparing high and low SGA–EPA Similarity, $t(210) = –.51$, *ns*. However, when the EPA came from a source low in credibility, participants placed significantly less weight on the EPA when SGA–EPA Similarity was high than when it was low, $t(210) = 8.04$, $p < .001$.

Confidence in final responses. The confidence data were analysed using a $2 \times 2 \times 2$ repeated-measures analysis of variance. Consistent with predictions, a main effect of SGA–EPA Similarity emerged, $F(1, 210) = 9.11$, $p < .01$, such that participants were more confident in their final responses when SGA–EPA Similarity was high ($M = 4.27$, $SD = 1.09$) than when it was low ($M = 3.88$, $SD = 1.21$). Also, as expected, a main effect of Source Credibility emerged, $F(1, 210) = 99.54$, $p < .001$; participants were more confident in their final responses when Source Credibility was high ($M = 4.71$, $SD = 1.02$) than when it was low ($M = 3.44$, $SD = 1.28$). The main effects were qualified by a statistically significant

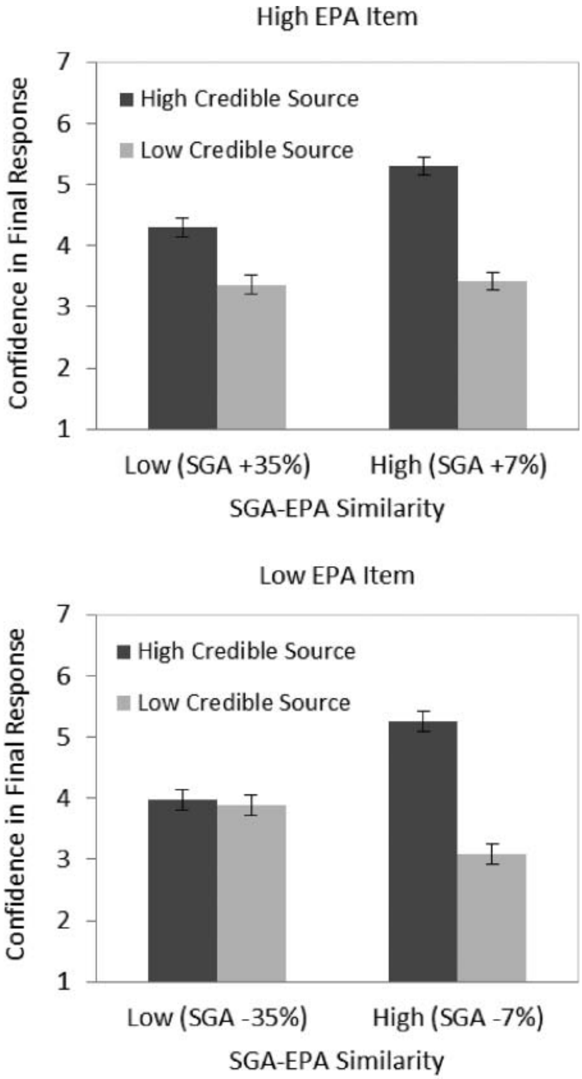

Figure 2. Mean confidence in final responses by EPA item, SGA–EPA similarity, and source credibility.

SGA–EPA Similarity × Source Credibility interaction, $F(1, 210) = 35.30$, $p < .001$. However, this two-way interaction was further qualified by an unexpected three-way interaction, $F(1, 210) = 8.96$, $p < .01$ (see Figure 2).

High EPA data. Among the high EPA data, a significant main effect of SGA–EPA Similarity emerged, $F(1, 210) = 11.85$, $p < .01$, such that greater confidence in final responses was reported when SGA–EPA Similarity was high ($M = 4.36$, $SD = 1.12$) than when it was low ($M = 3.83$, $SD = 1.11$). Additionally, a significant main effect of Source Credibility also emerged, $F(1, 210) = 85.56$, $p < .001$), such that greater confidence in final responses was reported when Source Credibility was high ($M = 4.80$, $SD = 1.12$) than when it was low ($M = 3.39$, $SD = 1.11$). Both of these main effects, however, were qualified by a significant SGA–EPA Similarity × Source Credibility interaction, $F(1, 210) = 5.23$, $p < .05$.

As expected, when SGA–EPA Similarity was high, significantly greater confidence in final responses was reported when the EPA came from a highly credible source, than when it came from a low-credible source, $t(210) = 10.13$, $p < .001$. This same general pattern was found when SGA–EPA Similarity was low, $t(210) = 5.04$, $p < .001$. Also consistent with expectations, when the EPA came from a highly credible source, significantly greater confidence in final responses was reported when SGA–EPA Similarity was high, than when it was low, $t(210) = 5.37$, $p < .001$. However, when the EPA came from a low credible source, there was no significant difference in reported confidence in final responses between high SGA–EPA Similarity and low SGA–EPA Similarity, $t(210) = 0.27$, *ns*.

Low EPA data. Among the low EPA data, only a significant main effect of Source Credibility emerged, $F(1, 210) = 47.31$, $p < .001$, such that greater confidence was reported when SGA–EPA Similarity was high ($M = 4.36$, $SD = 1.22$) than when it was low ($M = 3.49$, $SD = 1.21$). However, the main effect was qualified by a significant SGA–EPA Similarity × Source Credibility interaction, $F(1, 210) = 6.61$, $p < .05$.

As expected, when SGA–EPA Similarity was high, significantly greater confidence was reported when the EPA came from a highly credible source than when it came from a low-credible source, $t(210) = 11.65$, $p < .001$. However, when SGA–EPA Similarity was low, there was no significant difference in reported confidence when the EPA came from a highly credible source or low-credible source, $t(210) = 0.48$, *ns*. Also consistent with expectations, when the EPA came from a highly credible source, significantly greater confidence in final responses was reported when SGA–EPA Similarity was high than when it was low, $t(210) = 6.89$, $p < .001$. Unexpectedly, when the EPA came from a low-credible source, significantly greater confidence was reported when SGA–EPA Similarity was low than when it was high, $t(210) = 4.28$, $p < .001$.

It is worth noting that participants' ratings of confidence were greatest in the expected conditions. When the EPA was high, a planned contrast showed that confidence among participants in the high SGA–EPA

Similarity/high Source Credibility condition ($M = 5.30$, $SD = .89$) was greater than that for participants in all other conditions ($M = 3.69$, $SD = 1.18$), $t(210) = 10.59$, $p < .001$. This same result was found when the EPA was low; $M = 5.26$, $SD = .96$ vs $M = 3.65$, $SD = 1.27$, $t(210) = 10.58$, $p < .001$.

It is worth noting that participants' ratings of confidence were greatest when our metacognitive account predicted that they would be: the difference between high SGA–EPA Similarity/high Source Credibility when the EPA was high ($M = 3{,}406$, $SE = 37.61$) and high SGA–EPA Similarity/high Source Credibility when the EPA was low ($M = 1{,}984$, $SE = 37.37$) was highly significant, $t(207) = 26.79$, $p < .001$. Essentially, participants who gave the greatest final responses to the high EPA item were the same participants who gave the smallest final responses to the low EPA item; yet they were also the most confident participants (recall that the correct answers to both items were identical).

GENERAL DISCUSSION

Our findings suggest that when people encounter traditional anchoring items they activate SGAs, think about the implications of SGA–EPA similarity, remain sensitive to EPA source information, and to some degree integrate the EPA with their SGA to render a final response. We believe that this is a metacognitive process (see Jost, Kruglanksi, & Nelson, 1998), and that our experiment provides important evidence in favour of a metacognitive account of the anchoring heuristic. Clearly our results with regard to confidence in one's final response are important to our contention. Similar results may also be predicted from/supported by the attitudinal perspective of anchoring effects (Blankenship, Wegener, Petty, Detweiler-Bedell, & Macy, 2008; Wegener et al., 2010, 2001). However, no other existing research appears to measure SGAs and make explicit predictions about different degrees of confidence depending on the level of SGA–EPA similarity and EPA source credibility; we found these variables to interact to influence the degree of confidence that our participants expressed for their final responses.

We contend that our data reveal something novel about the anchoring process in the context of Mussweiler and Strack's (1999) notion of mentally testing the EPA as a hypothesis. Mussweiler and Strack's model suggests that once people are exposed to an EPA they tend to generate anchor-consistent knowledge which is then used to generate a subsequent absolute judgement. From Mussweiler and Strack's model one might predict that people will always have some reason to be confident in their final responses to anchoring questions. However, from our metacognitive account we predicted and found our participants to differ in how confident they were in their final responses. Only when the EPA was similar to their SGA, and when the EPA came from a credible source, did participants express substantial confidence in their final responses. This finding suggests that people do think about different aspects of the information they receive and the implications they have for their judgements. This, by definition, is a metacognitive process.

Importantly, our experiment details not only the degree to which people integrate the EPA into their final response (with respect to their SGA), but we have also detailed the conditions under which the EPA will be weighted. For instance, for both the high and low EPA items, we found participants to place greater weight on the EPA when the EPA came from a highly credible source than a low-credible source. Initially we expected the degree of SGA–EPA Similarity to influence the weight placed on the EPA in the final response when Source Credibility was high. However, when Source Credibility was high, participants placed substantial weight on the EPA regardless of whether SGA–EPA Similarity was high or low. This pattern of data is suggestive of an assimilation effect (i.e., when judgements are displaced towards context stimuli; see Mussweiler, 2007). We suspect that participants found source credibility relatively more important than SGA–EPA Similarity and considered it reasonable to assimilate their final responses towards the EPA when source credibility was considered high (irrespective of SGA–EPA Similarity). However, this does not rule out the possibility that participants recognised the implications of high SGA–EPA Similarity with a highly credible source. As we know from the work of Damisch, Mussweiler, and Plessner (2006), focusing on similarities between two stimuli tends to produce an assimilation effect, and attention to high SGA–EPA Similarity would seem especially reaffirming when the EPA comes from a highly credible source. In fact, evidence of our metacognitive account is provided by the fact that greater confidence was expressed in final responses in the context of high SGA–EPA Similarity/high Source Credibility compared to low SGA–EPA Similarity/high Source Credibility.

Perhaps our most intriguing finding was the fact that the degree of SGA–EPA Similarity influenced the weight placed on the EPA in the final response when Source Credibility was low; suggestive of a strong contrast effect (i.e., when judgements are displaced away from context stimuli; see Mussweiler, 2007). That is, participants placed less weight on the EPA when SGA–EPA Similarity was high than when it was low, and especially when the EPA came from a source low in credibility. In fact, the average weight placed on the EPA was negative when SGA–EPA Similarity was high and the EPA came from a low credible source. When a negative weight is obtained, it indicates that one's final response was influenced by the EPA but in the opposite direction of a compromise between his/her SGA and the EPA. Thus when one generates an SGA similar to an estimate originating from a source low in credibility, he/she may interpret this information as a cue that his/her

SGA may be incorrect as our participants polarised their final judgements even further away from the EPA in such cases. It is important to note that such cases did not result in as much confidence as that of their counterparts who were presented with the same EPA from an alleged highly credible source.

The present findings build on previous anchoring studies (Epley & Gilovich, 2001, 2004, 2005, 2006) that have found SGAs to be important for the anchoring effect. Specifically, data from the current research suggest that SGAs appear to activate and moderate processes of information integration such that people use their relevant, self-generated knowledge in a meaningful way when encountering anchoring items. We suggest that people appropriately factor in the EPA and its source to render a final response, a process conceptually similar to differentiating between relevant and irrelevant consensus information (see Gigerenzer, 2008; Gigerenzer & Goldstein, 1999).

A potential concern with the proposed metacognitive model is the question of whether or not SGAs are spontaneously generated or used only when they are made salient by an experimental paradigm. Consistent with prior research (e.g., Epley & Gilovich, 2001), the results of our pilot tests suggest that SGAs are spontaneously generated, and furthermore, simply requesting people to report their SGA does not influence their final responses.

While the current research addresses questions related to the process underlying anchoring, some questions remain. Our data demonstrated that metacognitively relevant information was used by participants to re-evaluate their initial estimate in the presence of externally salient information, which subsequently influenced their final responses. Given this finding, it seems likely that detailed information about an EPA may be used to justify the usefulness of the EPA in a metacognitive way. The more detail provided about an EPA the easier it may be to estimate its validity. In contrast, insufficient detail about the EPA should reduce the influence of the EPA because it might be perceived as invalid. Future research would do well to identify other source-based factors that influence information integration concerning the traditional anchoring paradigm.

It is worth noting that in the current research we employed relevant anchors, while many other anchoring studies tend to use irrelevant anchors (e.g., numbers that are supposedly randomly determined). Because our data suggest that people do attend to information relevant to the EPA, it seems possible that SGA–EPA similarity might only have an effect when an EPA is considered relevant. It is also likely that irrelevant EPAs may not result in the same degree of confidence as that of relevant EPAs.[2] Thus investigation with regard to the relevance of an EPA is warranted.

[2] We are indebted to an anonymous reviewer for pointing out the potential importance of this distinction.

Finally, the current perspective of anchoring warrants further investigation. Subsequent studies may further this line of work by borrowing from models of metacognition (e.g., Petty & Briñol, 2008). For example, investigating whether or not the effects shown here are augmented among high need for cognition individuals, conditions of high cognitive elaboration, or high EPA-argument quality are three cases whereby one might expect to find greater metacognitive processing.

Manuscript received 5 April 2012
Revised manuscript received 21 May 2013
First published online 12 July 2013

REFERENCES

Anderson, N. (1970). Functional measurement and psycho-physical judgement. *Psychological Review, 77*, 153–170.

Anderson, N. (1981). *Foundations of information integration theory*. New York, NY: Academic Press.

Anderson, N. (1982). *Methods of information integration theory*. New York, NY: Academic Press.

Blankenship, K. L., Wegener, D. T., Petty, R. E., Detweiler-Bedell, B., & Macy, C. L. (2008). Elaboration and consequences of anchored estimates: An attitudinal perspective on numerical anchoring. *Journal of Experimental Social Psychology, 44*, 1465–1476.

Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes, 79*, 115–153.

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgements of belief and value. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgement* (pp. 120–138). Cambridge, UK: Cambridge University Press.

Damisch, L., Mussweiler, T., & Plessner, H. (2006). Olympic medals as fruits of comparison? Assimilation and contrast in sequential performance judgements. *Journal of Experimental Psychology: Applied, 12*, 166–178.

Epley, N., & Gilovich, T. (2001). Putting the adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science, 12*, 391–396.

Epley, N., & Gilovich, T. (2004). Are adjustments insufficient? *Personality and Social Psychology Bulletin, 30*, 447–460.

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgemental anchoring: Differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision-Making, 87*, 327–339.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17*, 311–318.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science, 3*, 20–29.

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In G. Gigerenzer, P.M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 75–95). New York, NY: Oxford University Press.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin, 21*, 1161–1166.

Jarvis, W. B. G. (2006). *MediaLab* v2006 [computer software]. New York, NY: Empirisoft.
Jost, J. T., Kruglanksi, A. W., & Nelson, T. O. (1998). Social metacognition: An expansionist review, *Personality and Social Psychology Review, 2*, 137–154.
Matthews, W. J. (2011). What might judgement and decision making research be like if we took a Bayesian approach to hypothesis testing? *Judgement and Decision Making, 6*, 843–856.
McElroy, T., & Dowd, K. (2007). Susceptibility to anchoring effects: How openness-to-experience influences responses to anchoring cues. *Judgement and Decision Making, 2*, 48–53.
Mussweiler, T. (2007). Assimilation and contrast as comparison effects: A selective accessibility model. In D. A. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology* (pp. 165–185). New York, NY: Psychology Press.
Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 35*, 136–164.
Mussweiler, T., & Strack, F. (2001). Considering the impossible: Explaining the effects of implausible anchors. *Social Cognition, 19*, 145–160.
Petrocelli, J. V., Tormala, Z. L., & Rucker, D. D. (2007). Unpacking attitude certainty: Attitude clarity and attitude correctness. *Journal of Personality and Social Psychology, 92*, 30–41.
Petty, R. E., & Briñol, P. (2008). Persuasion from single to multiple to metacognition. *Perspectives on Psychological Science, 3*, 137–147.
Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163.
Ritov, I. (1996). Anchoring in simulated competitive market negotiation. *Organisational Behavior and Human Decision Processes, 67*, 16–25.
Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review, 16*, 225–237.
Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
Wegener, D. T., Petty, R. E., Blankenship, K. L., & Detweiler-Bedell, B. (2010). Elaboration and numerical anchoring: Implications of attitude theories for consumer judgement and decision making. *Journal of Consumer Psychology, 20*, 5–16.
Wegener, D. T., Petty, R. E., Detweiler-Bedell, B. T., & Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology, 37*, 62–69.
Yaniv, I. (1997). Weighting and trimming: Heuristics for aggregating judgements under uncertainty. *Organisational Behavior and Human Decision Processes, 69*, 237–249.
Yaniv, I. (2004a). Receiving other people's advice: Influence and benefit. *Organisational Behavior and Human Decision Processes, 93*, 1–13.
Yaniv, I. (2004b). The benefit of additional opinions. *Current Directions in Psychological Science, 13*, 75–78.
Yaniv, I., & Foster, D. P. (1997). Precision and accuracy of judgemental estimation. *Journal of Behavioral Decision Making, 10*, 21–32.
Yaniv, I., & Milyavsky, M. (2007). Using advice from multiple sources to revise and improve judgements. *Organisational Behavior and Human Decision Processes, 103*, 104–120.