Organizational Behavior and Human Decision Processes 118 (2012) 216–225

Contents lists available at SciVerse ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

Self-selection bias in hypothesis comparison

Jennifer C. Whitman, Todd S. Woodward *

Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada
BC Mental Health and Addictions Research Institute, University of British Columbia, Vancouver, BC, Canada

ARTICLE INFO

Article history:
Received 18 December 2010
Accepted 21 February 2012
Available online 20 March 2012
Accepted by Julie Irwin

Keywords:
Judgment
Probabilistic reasoning
Decision-making
Overconfidence
Choice
Option fixation
Hypothesis generation

ABSTRACT

Here we investigated whether, given equivalent supporting evidence, we judge self-selected hypotheses differently from those selected by an external source. On each trial of a probabilistic reasoning task requiring no retrieval from memory, participants rated the probability of a focal hypothesis, relative to two alternatives. The focal hypothesis was either selected by the participant or by a computer. In four experiments, self-selected focal hypotheses were judged to be more probable than externally selected ones, despite equivalent supporting evidence. This self-selection bias was independent of level of difficulty in selecting the focal hypothesis (cognitive effort) and of whether evidence was gradually accumulated or all presented instantaneously. These results suggest that the cognitive operations involved in selecting a hypothesis lead to assignment of higher probability to that hypothesis, and that this effect is independent of hypothesis selection difficulty and of the rate of evidence accumulation.

Introduction

When attempting to understand a situation, we often generate or select our own preferred hypothesis, and consider it alongside those provided by an external source such as a news report, an advertisement, or someone else's opinion. Forming our own opinion often involves assessing our level of agreement with all candidate hypotheses. The main goal of our series of experiments was to investigate whether judgments can be biased in favor of self-selected hypotheses, relative to judgments of hypotheses specified by external sources, given equivalent supporting evidence. The secondary goals were to test whether or not this effect was moderated by other factors relating to hypothesis selection difficulty and the rate of evidence accumulation.

Previous investigations involving comparison of self-selected to other-selected material have required participants to choose from a range of options and rate their confidence that they chose correctly. This is compared to confidence that other-selected material was chosen correctly (Koehler, 1994; Ronis & Yates, 1987; Sieck, 2003; Sieck, Merkle, & Van Zandt, 2007; Sieck & Yates, 2001; Sniezek, Paese, & Switzer, 1990). In the general knowledge and learning paradigms used in those studies, self-selection and confidence ratings are likely affected by familiarity with the subject matter, fluency in memory retrieval, or completeness of memory retrieval (Sieck, 2003; Sieck et al., 2007; Sieck & Yates, 2001). As the experimenter has no control over these processes, it is difficult (or impossible) to equate them between the self-selected and other-selected conditions. Furthermore, in such an experimental setup, the assignment of a given question to the self-selected vs. other-selected conditions was necessarily manipulated between-subjects. Consequently, the evidence considered for that question likely differed between the individual making the ratings in the self-selected condition and the individual making the ratings in the other-selected condition. Various aspects of familiarity, personal experience, fluency, and completeness of memory retrieval for a question may differ between the individual who chooses one answer before rating it and the other individual who rates their level of agreement with that choice, making it difficult to isolate the factors causing ratings to differ between the self-selected and other-selected conditions.

In the current research, we wished to precisely match supporting evidence across the self-selected and other-selected conditions and avoid any differences in familiarity, past personal experience, or retrieval. To this purpose, we used a variation of the well-known beads-from-a-jar probabilistic reasoning task (Beach, 1968; Freeman, Pugh, & Garety, 2008; Moritz, Woodward, & Lambert, 2007; Peterson, Schneider, & Miller, 1965; Speechley, Whitman, & Woodward, 2010; Whitman & Woodward, 2011). This task involves judging the likelihood that a series of beads is drawn from jar A and not jar B, based on the colors of the beads in jars A and B, and on the colors in the series of beads being drawn. Given that jars A and B differ only in their relative proportions of bead colors, the process of self-selecting can be based on a very restricted set of parameters, presumably almost entirely probability estimates. This paradigm

* Corresponding author. Address: Room A3-A116, BC Mental Health & Addictions Research Institute – Translational Research Building, 3rd Floor, 938W. 28th Avenue, Vancouver, British Columbia, Canada V5Z 4H4. Fax: +1 604 875 3871.
E-mail address: Todd.S.Woodward@gmail.com (T.S. Woodward).

0749-5978/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.obhdp.2012.02.002

allows us to equate evidence within-subjects between the self-selected and other-selected conditions. The probability ratings following self- or other-selection of hypotheses can then be compared.

The version of the task used in the current study involved judging the probability that a single item was drawn from one of three containers. Specifically, participants judged the probability that the item was drawn from a given self- or other-selected container rather than being drawn from either of two alternatives. The self- or other-selected container is referred to as the *focal hypothesis*. On all trials, the participant was required to indicate the likelihood of the focal hypothesis (rather than the other two alternatives) being true. We predicted a *self-selection bias* whereby ratings of the relative probability of the focal hypothesis would be higher if it were self-selected than if it were externally selected. Our secondary goal was to explore possible moderators of any self-selection bias, such as increased cognitive effort inherent to the selection process, or repeated activation of the mental representation of the self-selected focal hypothesis.

Experiment 1

The purpose of Experiment 1 was to assess whether self-selected hypotheses were judged to be more probable than externally selected ones, despite their being matched in terms of the mathematically normative probability rating. Our task was designed so that any such bias could not be attributable to either differences in the number of alternatives considered or differences in memory retrieval.

Method

Participants

Twenty-six volunteers (14 females and 12 males) with a mean age of 29.3 years ($SD = 9.2$) participated in this experiment. Participants were recruited via posters on the University of British Columbia campus and in community centers in the greater Vancouver area, and also via postings on electronic bulletin boards. All participants were reimbursed $10 per hour for their time plus parking and transportation expenses.

Materials and procedure

Each trial of our probabilistic reasoning task involved a scene depicting four blue lakes (see Fig. 1), three of which were upstream from the fourth. At the start of each trial, a single black or white fish was seen to jump from the downstream lake. The color of this jumping fish, referred to hereafter as the relevant color, was randomized across trials. Subsequently, the contents of each of the three upstream lakes became visible. Each contained a mixture of black and white fish (40 fish in total per lake), which remained in view until a rating had been made. Participants were told that the fish in the downstream lake originated from one of the upstream lakes. Thus, each upstream lake corresponded to a hypothesis about the origin of the jumping fish. The upstream lake with the most fish of the relevant color was the most likely origin of the fish in the downstream lake.

Subsequently, if the trial was one in which the focal hypothesis was to be self-selected, participants indicated the lake they deemed to be the most likely origin of the jumping fish by moving the mouse cursor (a red square) over their preferred lake and clicking with the left mouse button. This lake, referred to hereafter as the focal lake, then turned green. On trials with externally selected focal hypotheses, the focal lake was selected randomly from the three choices by the computer, and turned green immediately after

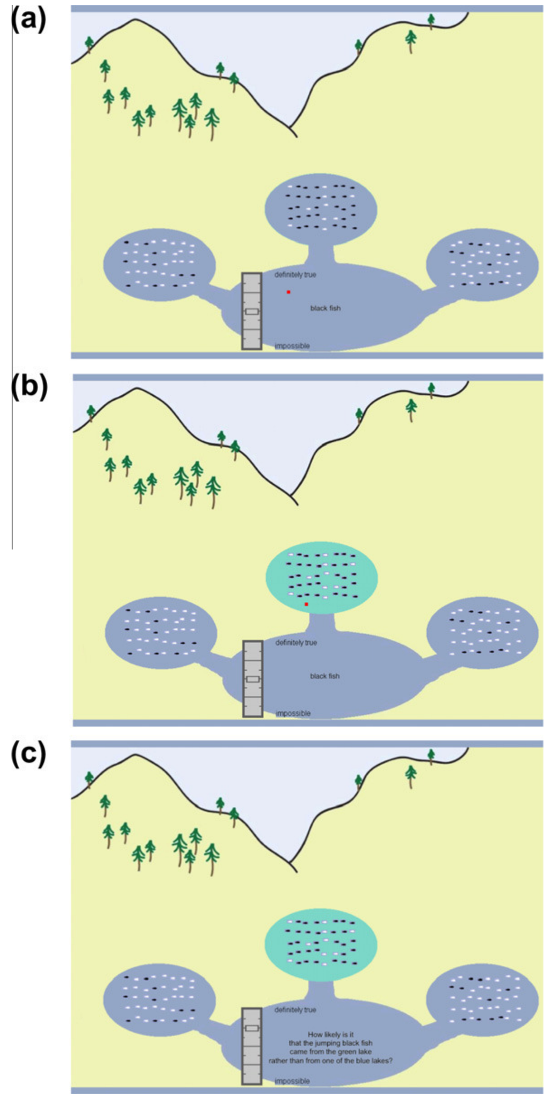

Fig. 1. Sequence of events within a trial of either Experiment 1, 2, or 3 in which the focal hypothesis is self-selected. The trial starts with a single fish, either black or white, jumping from the downstream lake. Participants are informed that the jumping fish originated in one of the three upstream lakes. (a) After the fish jump, the text next to the rating scale reminds the participant of its color. (b) When the participant clicks on the lake considered to be the most likely origin (selects the focal hypothesis), it turns green. (c) The participant then rates the probability that the jumping fish came from the green focal lake rather one of the two blue alternate lakes by moving a cursor up or down the rating scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the fish in the bottom lake jumped, at the same time as the fish in the upstream lakes became visible.

Once the focal lake was selected, participants then rated the relative probability that the focal hypothesis, rather than its alternative, was true. Ratings were made on a vertical scale, with the labels "definitely true" at the top end and "impossible" at the bottom end. Participants used a mouse to move the slider smoothly up and down the scale and clicked the left mouse button when it was in the desired location. The vertical position of the cursor, in pixels, was then recorded. In the results section, ratings are reported as a percentage of the total height of the response scale. At the beginning of each trial, the slider was set at the mid-point of the response scale. Trials with self-selected and externally selected focal hypotheses occurred in separate blocks. In total there were 18 blocks of nine trials each, alternating between blocks in which the focal hypothesis was self-selected and blocks in which it was externally selected. The first block always consisted of trials in which the focal hypothesis was externally selected (by the computer).

In order to vary the strength of support for the focal hypothesis, as well as the strength of support for its alternatives, we manipulated the percentage of fish of the relevant color in each lake. Specifically, each of the three lakes contained either 25%, 50%, or 75% fish of the relevant color. Any trials on which the participant made an error in selecting the most probable hypothesis as the focal hypothesis were rejected from the analysis. Note that this prevented us from employing a $3 \times 3 \times 3$ design (3 levels of support for each lake), because support for the focal hypothesis had to match or exceed support for each alternative. Thus, within a given type of focal hypothesis (self-selected or externally selected), we used a design with 10 cells, as displayed in Table 1.

Results and discussion

In order to ensure that participants clearly understood the instructions, we excluded their data if their accuracy in selecting the most probable hypothesis as the focal hypothesis was less than 80%. As a result, six participants were excluded from the analysis. For the remaining 20 participants, we excluded any trials in which the focal hypothesis was less probable than one of its alternatives, whether that focal hypothesis was selected by the participant or computer-selected. We excluded an average of 3.3% (*SD* = 3.9%) of trials with self-selected focal hypotheses. As the computer was randomly selecting focal hypotheses, and as 19 of the 30 choices outlined in the first three columns of Table 1 are correct choices, the computer's accuracy in this experiment was 63.3%. Consequently, we excluded 36.7% of the trials with computer-selected focal hypotheses. Mean ratings for each condition are displayed in Table 1.

A *t*-test comparing the average rating of self-selected focal hypotheses to the average rating of externally selected ones revealed a significant self-selection bias, $t(19) = 4.32$, $p < .001$, $\eta^2 = .50$, involving higher mean probability ratings for self-selected focal hypotheses ($M = 59.4$) than for externally selected ones ($M = 56.2$). In order to adjust for moderate skew in the data, we also applied a rank transformation (Conover & Iman, 1981). Specifically, the entire set of scores (both dependent variables together) was rank-transformed, then the *t*-test was applied to the rank-transformed scores. When these data were transformed to rank scores, the effect remained significant. The ratings of self-selected focal hypotheses were higher for 16 of the 20 participants. Unlike in previous paradigms involving judgments of one's own hypothesis (Koehler, 1994; Ronis & Yates, 1987; Sieck, 2003; Sieck et al., 2007; Sieck & Yates, 2001; Sniezek et al., 1990), the bias reported here cannot be attributed to differences in familiarity, personal experience, or memory retrieval. It is consistent with the notion that the cognitive operations involved in selecting one's own preferred focal hypothesis increase its perceived probability.

As can be seen in Table 1, the average rating, across conditions, is higher than the normative average. This may involve neglect of the fact that the base rate was 1/3, which would be consistent with previous reports of base rate neglect (Tversky & Kahneman, 1974). It may also be a bias induced by the demand characteristics of the experiment, because at the start of each trial the cursor was at the midpoint of the rating scale, rather than 1/3 of the way from the bottom. Regardless, both the base rate and the starting position of the mouse cursor on the rating scale are equated between the self-selected and externally selected conditions. Consequently, this overall tendency to make ratings higher than the norm is irrelevant to the main theme of this paper, namely the selection bias.

Experiment 2

One possible explanation for the results of Experiment 1 was that participants correctly noticed that the computer was sometimes selecting something other than the most probable hypothesis to be the focal hypothesis, and thus considered the computer's choices to be random or inaccurate. The purpose of Experiment 2 was to replicate the results of Experiment 1 while matching the computer's choices to the participant's. This would ensure that the self-selection bias was not just the result of the participant generally assigning low ratings to the computer-selected focal hypotheses due to a perception of the computer as unreliable in selecting the focal hypothesis correctly.

Method

Participants

Forty-one volunteers (21 females and 20 males) with a mean age of 27.7 years (*SD* = 8.5) participated in this experiment. Participants were recruited via posters on the University of British Columbia campus and in community centers in the greater Vancouver area, and also via postings on electronic bulletin boards. All participants were reimbursed $10 per hour for their time plus parking and transportation expenses.

Table 1
Experiment 1: ratings of the relative probability that the focal hypothesis, rather than its alternatives, is true, as a function of the strength of the evidence supporting the focal hypothesis and of that supporting each alternative.

Strength of evidence supporting each hypothesis (% fish in that lake of relevant color)				Rating (percentage of response scale height)		Self-selection bias
Focal hypothesis (*F*)	Alternative #1 (*A*1)	Alternative # 2 (*A*2)	Mathematical norm $100 * F/(F + A1 + A2)$	Self-selected focal	Externally selected focal	
75	75	75	33.33	53.95	48.68	5.27
75	50	75	37.50	58.82	52.98	5.84
75	50	50	42.86	57.81	57.22	0.59
75	25	75	42.86	63.11	57.83	5.27
75	25	50	50.00	61.82	61.53	0.29
75	25	25	60.00	68.31	66.04	2.27
50	50	50	33.33	50.62	45.79	4.83
50	25	50	40.00	54.38	49.94	4.44
50	25	25	50.00	57.40	55.32	2.08
25	25	25	33.33	42.52	42.48	0.04

Note: Each hypothesis corresponds to an upstream lake which is the possible origin of a fish seen in a downstream lake. The strength of evidence supporting each corresponds to the percentage of fish in that lake matching the color of the downstream fish.

Table 2
Experiment 2: ratings of the relative probability that the focal hypothesis, rather than its alternatives, is true, as a function of the strength of the evidence supporting the focal hypothesis and of that supporting each alternative.

Strength of evidence supporting each hypothesis (% fish in that lake of relevant color)				Rating (percentage of response scale height)		Self-selection bias
Focal hypothesis (*F*)	Alternative #1 (*A*1)	Alternative # 2 (*A*2)	Mathematical norm 100 * *F*/(F + *A*1 + *A*2)	Self-selected focal	Externally selected focal	
75	75	75	33.33	42.15	41.94	0.21
75	50	75	37.50	51.90	49.18	2.72
75	50	50	42.86	58.33	54.25	4.08
75	25	75	42.86	55.10	51.93	3.17
75	25	50	50.00	62.60	60.72	1.88
75	25	25	60.00	68.50	67.93	0.57
50	50	50	33.33	43.50	40.81	2.69
50	25	50	40.00	51.37	48.66	2.71
50	25	25	50.00	58.46	55.23	3.23
25	25	25	33.33	39.88	40.30	−0.42

Materials and procedure

The methods of Experiment 2 were identical to those of Experiment 1 with the following exceptions. The first block always consisted of trials in which the focal hypothesis was self-selected. Each trial with an externally selected focal hypothesis was essentially a repeat of a corresponding trial involving a self-selected focal hypothesis in the preceding block. Specifically, the choice of focal lake (focal hypothesis) made by the computer always matched a choice made by the participant on the corresponding trial in the preceding block. Thus, the computer's accuracy in selecting the most probable hypothesis was identical to the participant's accuracy. Within a matched pair of trials, the proportions of fish of each color within each lake were kept constant. However, the positions of individual fish within those lakes varied randomly.

Results and discussion

In order to ensure that participants clearly understood the instructions, we excluded their data if their accuracy in selecting the most probable hypothesis as the focal hypothesis was less than 80%. As a result, two participants were excluded from the analysis. For the remaining 39 participants, we excluded any trials with selection errors (trials in which the self-selected focal hypothesis was less probable than one of its alternatives), whether that selection was made by the participant or by the computer. We excluded an average of 2.8% ($SD = 3.5\%$) of trials with self-selected focal hypotheses. The percentages were the same for trials with externally selected focal hypotheses, as accuracy was matched between the self and external conditions. Mean ratings for each condition are displayed in Table 2. A *t*-test comparing the average rating of self-selected focal hypotheses to the average rating of externally selected ones revealed a significant self-selection bias, $t(38) = 3.75$, $p < .001$, $\eta^2 = .27$, involving higher mean probability ratings for self-selected focal hypotheses ($M = 53.3$) than for externally selected ones ($M = 51.6$). As in Experiment 1, the effect remained significant when the data were transformed to rank scores. The ratings of self-selected focal hypotheses were higher for 30 of the 39 participants. Thus, the results of Experiment 2 replicated those of Experiment 1 and showed that the self-selection bias could not be attributed to overall differences in accuracy, between the participant and the computer, in selecting the most probable focal hypothesis. The smaller effect size in Experiment 2 than Experiment 1 is consistent with our concerns that perceptions of the computer's accuracy contributed in part to the selection bias in Experiment 1. This highlights the importance of the self-to-external accuracy matching implemented here and in our subsequent experiments.

Experiment 3

The process of selecting a focal hypothesis and then rating its probability (relative to its alternatives) might involve more total cognitive effort than merely judging the probability of an externally selected focal hypothesis. Thus, selecting a focal hypothesis might cause the salience of its mental representation to be greater than that of an externally selected focal hypothesis via more effortful processing. If so, then we would expect the strength of the self-selection bias to depend on the amount of cognitive effort involved in selecting the focal hypothesis. The goals of Experiment 3 were to replicate the self-selection bias found in Experiments 1 and 2, and to determine whether that bias depended on increased cognitive effort involved in focal hypothesis selection.

The process of selecting a preferred focal hypothesis with consistent accuracy necessarily involves comparisons between evidence supporting competing hypotheses. Comparisons made between similar numbers are more difficult than comparisons made between dissimilar ones (Pinel, Dehaene, Riviere, & LeBihan, 2001). Thus, selecting the most probable hypothesis as the focal hypothesis will be most difficult if the strength of its supporting evidence is similar to the strength of the evidence supporting its alternatives. In the current experiment, we manipulated the degree of similarity between support for the focal hypothesis and support for its alternatives. The strength of support for the strongest alternative was either identical to that for the focal hypothesis, close (slightly lower), distant, or very distant. The support for the other alternative was varied in the same manner. If self-selection bias were the result of increased cognitive effort, we would expect it to be largest when support for the alternatives was most similar to support for the focal hypothesis.

Method

Participants

Seventy-nine volunteers (53 females and 26 males) with a mean age of 27.7 years ($SD = 7.9$) participated in this experiment. Participants were recruited via posters on the University of British Columbia campus and in community centers in the greater Vancouver area, and also via postings on electronic bulletin boards. All participants were reimbursed $10 per hour for their time plus parking and transportation expenses.

Materials and procedure

The methods were identical to those of Experiment 2 with the following exceptions. In total there were twelve blocks of ten trials each, alternating between blocks in which the focal hypothesis was self-selected and blocks in which it was externally selected. The fo-

Table 3
Experiment 3: ratings of the relative probability that the focal hypothesis, rather than its alternatives, is true, as a function of the strength of the evidence supporting the focal hypothesis and of that supporting each alternative.

Strength of Evidence			Math.	Rating		Self-selection bias (self-ext)
Focal	Alt. 1	Alt. 2	Norm	Self	External	
85	85	85	33.33	44.86	42.89	1.97
85	85	75	34.69	48.44	45.88	2.56
85	75	75	36.17	50.12	48.10	2.02
85	85	45	39.53	54.69	52.88	1.81
85	85	15	45.95	56.91	56.37	0.54
85	75	45	41.46	55.51	53.49	2.02
85	75	15	48.57	59.99	58.64	1.35
85	45	45	48.57	62.39	60.47	1.92
85	45	15	58.62	65.85	65.64	0.21
85	15	15	73.91	75.21	74.28	0.93
55	55	55	33.33	45.12	44.17	0.95
55	55	45	35.48	45.60	44.78	0.82
55	45	45	37.93	46.36	45.94	0.42
55	55	15	44.00	53.30	53.52	−0.22
55	45	15	47.83	54.76	52.59	2.17
55	15	15	64.71	63.79	63.23	0.56
25	25	25	33.33	42.06	43.03	−0.97
25	25	15	38.46	45.34	45.68	−0.34
25	15	15	45.45	48.13	49.14	−1.01

cal lake contained either: 85%, 55%, or 25% fish of the relevant color. Support for each alternative hypothesis was either identical to the support for the focal hypothesis, close (10% lower), distant (40% lower), or very distant (70% lower). Thus, each alternative lake contained either 85%, 75%, 55%, 45%, 25%, or 15% fish of the relevant color. These manipulations are outlined in Table 3, which shows a complete list of the conditions used in this experiment. They allowed us to test whether the magnitude of any self-selection bias varied as a function of the difficulty of selecting the most probable hypothesis; in other words, the degree to which the focal and alternative hypotheses had similar levels of support.

Results and discussion

Unlike in Experiments 1 and 2, all of the participants in Experiment 3 had accuracy levels above 80% in selecting the most probable hypothesis as the focal hypothesis. Thus, all of the participants appeared to understand the instructions clearly, and none were excluded. We excluded any trials in which the focal hypothesis was less probable than one of its alternatives, whether it was selected by the participant or by the computer. We excluded an average of 2.1% ($SD = 3.5\%$) of trials with self-selected focal hypotheses. The percentages were the same for trials with externally selected focal hypotheses, as accuracy was matched between the self and external conditions. Mean ratings for each condition are displayed in Table 3.

Replication of self-selection bias

We first tested for a self-selection bias as a main effect across all conditions. A t-test comparing the average rating of self-selected focal hypotheses to the average rating of externally selected ones revealed a significant self-selection bias, $t(78) = 2.82$, $p < .01$, $\eta^2 = .09$. As in Experiments 1 and 2, this involved higher ratings for self-selected focal hypotheses ($M = 53.6\%$) than for externally selected ones ($M = 52.7\%$). The ratings of self-selected focal hypotheses were higher for 47 of the 79 participants. As in Experiments 1 and 2, the effect remained significant when the data were transformed to rank scores.

Table 4
Experiment 3: time taken to self-select a focal hypothesis and to make ratings of the relative probability that the focal hypothesis, rather than its alternatives, is true, as a function of the strength of the evidence supporting the focal hypothesis and of that supporting each alternative.

Strength of evidence			Time spent self-selecting focal hypothesis (ms)	Rating		Self-ext rating RT difference (ms)
Focal	Alt. 1	Alt. 2		Self	External	
85	85	85	6038.10	3225.43	7313.88	−4088.45
85	85	75	5543.75	3930.27	7494.03	−3563.76
85	75	75	4408.51	3585.57	7539.13	−3953.57
85	85	45	5329.80	3685.35	7706.17	−4020.82
85	85	15	5228.03	3781.71	7452.07	−3670.36
85	75	45	3917.59	4002.34	8085.35	−4083.01
85	75	15	3628.33	3835.74	7245.88	−3410.13
85	45	45	3373.96	4366.54	7705.75	−3339.21
85	45	15	3172.42	3763.80	6904.92	−3141.13
85	15	15	3205.72	3620.58	6829.95	−3209.37
55	55	55	6116.67	3082.15	8288.22	−5206.07
55	55	45	6220.15	3033.42	8657.94	−5624.52
55	45	45	5401.87	2933.10	7710.29	−4777.19
55	55	15	4931.23	3626.64	7979.03	−4352.40
55	45	15	4585.96	3723.67	8064.45	−4340.77
55	15	15	3529.96	4293.59	7473.59	−3179.99
25	25	25	6785.00	3273.57	8779.65	−5506.09
25	25	15	5421.04	3766.64	8669.46	−4902.82
25	15	15	4732.58	4234.24	7730.85	−3496.62

Similarity of focal hypothesis to alternatives

Next, we tested the prediction that self-selection bias would be strongest when it was most difficult to select the focal hypothesis because its alternatives had similar levels of support. We computed the magnitude of the self-selection bias separately for each participant and each combination of support for the focal hypothesis and support for the alternatives. This was used as the dependent variable in a one-way ANOVA. In order to investigate a wide range of Closeness of Alternatives, we constrained this analysis to the two highest levels of Support for the Focal Hypothesis (85% and 55% fish of the relevant color in the green lake). There were six levels of Closeness of Alternatives. In terms of how the percentage of fish of the relevant color in each alternate lake differed from that in the focal lake, the six levels of Closeness of Alternatives were (1) both identical to the focal lake, (2) one identical and one 10% lower, (3) both 10% lower, (4) one identical and one 40% lower, (5) one 10% lower and one 40% lower, and (6) both 40% lower. As a manipulation check, we assessed whether the time taken to select the focal hypothesis was longest for the conditions that were presumably most difficult, when the alternatives had almost as much support as the focal hypothesis. We found that selection speed did increase as Closeness of Alternatives increased, $F(5,390) = 18.85$, $p < .001$, $\eta^2 = .20$, as can be seen in Table 4. However, in the probability ratings there was no significant effect of Closeness of Alternatives on the magnitude of the self-selection bias, $F(5,390) = 0.47$, $p = .80$. This did not support the notion that self-selection bias depended on cognitive effort.

Experiment 4

When the focal hypothesis is self-selected, evidence consistent with it is processed twice: first during focal hypothesis selection, and again when making a judgment of relative probability. In contrast, that evidence is processed only once if the focal hypothesis was externally selected. Repeated processing of supporting evidence might cause the mental representation of the focal hypothesis to be more salient as a consequence of repeated activation. In this experiment we investigated this possibility using a manipulation that we developed in previous work with the same paradigm.

In previous work (Whitman & Woodward, 2011), we identified a bias towards giving more weight to evidence if it was accumulated gradually than if the same evidence was all presented instantaneously. Our account of this effect was that processing evidence to a greater extent would lead to more salient mental representations, and consequently greater subjective strength. The relevance of this to the current study was that this might share a common mechanism with the self-selection bias, relating to repeated activation of the mental representation of the focal hypothesis as a result of repeated processing of supporting evidence. If so, we would expect the selection bias and the overweighting of gradually accumulated evidence to be correlated across participants. In Experiment 4, we tested whether these two effects were correlated. We also considered the possibility that the selection bias might depend on whether evidence was gradually accumulated vs. all presented instantaneously. In order to be able to test the above possibilities, we orthogonally manipulated (1) whether the focal hypothesis was self-selected or externally selected and (2) whether evidence was gradually accumulated vs. instantaneously presented.

Our design included gradual evidence accumulation trials, in which only part of the evidence relevant to each hypothesis was presented on the first of two events. After an initial rating of relative probability was made, the remainder of the evidence relevant to each hypothesis was presented and a final revised rating was made. The gradual evidence accumulation trials were compared to instantaneous evidence control trials involving a single event on which all of the evidence relevant to each hypothesis was presented. On gradual evidence accumulation trials, the evidence obtained on the second event either increased the relative probability of the focal hypothesis, decreased it, or left it unchanged (the neutral condition).

Method

Participants

Forty-two volunteers (26 females and 16 males) with a mean age of 25.0 years ($SD = 7.6$) participated in this experiment. Participants were recruited via posters on the University of British Columbia campus and in community centers in the greater Vancouver area, and also via postings on electronic bulletin boards. All participants were reimbursed $10 per hour for their time plus parking and transportation expenses.

Materials and procedure

As in Experiments 1–3, the experiment alternated between blocks in which the focal hypothesis was self-selected and blocks in which it was externally selected (by the computer). The first block always consisted of trials in which the focal hypothesis was self-selected. In total, there were eight blocks of twelve trials each (four blocks in which the focal hypothesis was always self-selected, and four blocks in which it was externally selected).

The scene depicted on each trial of Experiment 4 was similar to that used in Experiments 1–3, except that the contents of each upstream lake were sometimes obscured near the outer edges, as shown in Fig. 2. Specifically, 25 of the 40 fish within each upstream lake were obscured on the first event of each trial with gradually accumulated evidence. On the second event of each of these trials, the entire contents of each lake became visible. On trials with instantaneously presented evidence, there was only one event, and the entire contents of each upstream lake were visible.

The sequence of events within a trial with gradually accumulated evidence and a self-selected focal hypothesis was as follows. First, a single black or white fish would be seen to jump from the downstream lake. Subsequently, the partially obscured contents of each of the three blue upstream lakes became visible. The participant indicated the lake most likely to be the origin of the downstream fish by moving the red mouse cursor over that lake and clicking on it. That lake then turned green. Subsequently, the participant rated the relative probability that the jumping fish came from the green upstream lake, rather than one of the blue upstream lakes, on the vertical rating scale. After that, the full contents of each upstream lake became visible, and the participant revised her or his rating by shifting the cursor on the response scale from the position corresponding to the previous rating to the position corresponding to the new rating. Within a trial with instantaneously presented evidence, the entire contents of each upstream lake became visible instantaneously after the downstream fish had jumped. Once the focal lake had been selected, a single rating of relative probability was made. For each of the above two types of trials with self-selected focal hypotheses (gradually presented evidence and instantaneous evidence), there were matching trials with externally selected focal hypotheses.

Within the four main trial types described above, there were four conditions defined by the proportions of fish in the upstream lakes. These are most easily described in terms of the sequence of events on trials with gradually accumulated evidence. On the first trial of each event, one of the three partially obscured upstream lakes would appear to contain 80% fish of the relevant color (12 of the 15 visible fish), another would appear to contain 40% fish of the relevant color (6 of the 15 visible fish), and the third lake would appear to contain 20% fish of the relevant color (3 of the 15 visible fish). The lake with 80% fish of the relevant color was

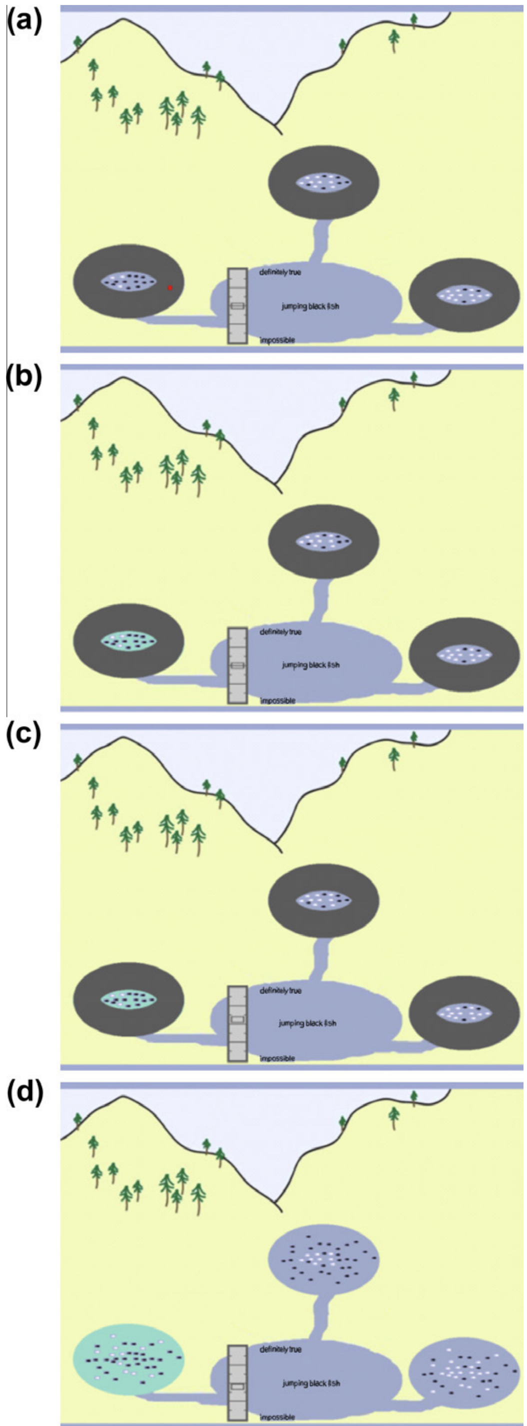

Fig. 2. Sequence of events within a trial of Experiment 4 in which the focal hypothesis is self-selected. At the beginning of the trial, each of the upstream lakes is partially occluded. Otherwise, the selection of the focal hypothesis (a and b) and subsequent rating (c) are the same as in Experiments 1–3. (d). Each of the upstream lakes then becomes visible in its entirety, and the participant revises the rating.

thus the obvious choice for the most probable location of the jumping fish (the focal hypothesis).

The evidence presented on the second event then either confirmed the previous choice of focal hypothesis, disconfirmed it, or was neutral. In the Confirmatory Evidence condition, the percentage of fish of the relevant color in the focal lake increased from 80% to 90%. The percentage in one alternative lake decreased from 40% to 15%. In the other it decreased from 20% to 10%. In the Disconfirmatory Evidence condition, the percentage of fish of the relevant color in the focal lake decreased from 80% to 70%. The percentage of fish in one alternative lake increased from 40% to 77.5%. In the other it increased from 20% to 47.5%. Recall from the previous three experiments (Table 1–3) that the relative probability of the focal hypothesis is defined as the strength of evidence supporting the focal hypothesis divided by the sum of the evidence supporting all three hypotheses. In the Confirmatory Evidence condition, this increased from 57% to 90/(90 + 15 + 10) = 78%. In the Disconfirmatory Evidence condition, it decreased from 57% to 70/(70 + 77.5 + 47.5) = 36%. Thus, the relative probability of the focal hypothesis either increased or decreased by 21 percentage points.

There were also two types of neutral evidence condition. The relative probability of the focal hypothesis stayed constant at 57% in both of these conditions. In the Neutral-Up condition, the percentage of fish of the relevant color in the focal lake increased from 80% to 90%, while the percentage in one alternative lake increased from 40% to 45% and the percentage in the other alternative lake increased from 20% to 22.5%. The relative probability of the focal hypothesis on the second event was thus 90/(90 + 45 + 22.5) = 57%. In the Neutral-Down condition, the percentage of fish of the relevant color in the focal lake decreased from 80% to 70%, while the percentage in one alternative lake decreased from 40% to 35% and the percentage in the other alternative lake decreased from 20% to 17.5%. The relative probability of the focal hypothesis on the second event was thus 70/(70 + 35 + 17.5) = 57%.

For each of the four conditions in which evidence was gradually accumulated (two events per trial), there was a corresponding condition in which it was all presented instantaneously (one event per trial). The evidence presented on an instantaneous-evidence trial was identical to the evidence on the second event of the corresponding gradually-accumulated-evidence trial. For example, for each confirmatory-evidence trial in the gradual evidence condition, there was a corresponding confirmatory-evidence-control trial in the instantaneous evidence condition. The evidence visible on the control trial was the same as the evidence visible on the final event of the gradual evidence trial. If, hypothetically, there were no biases caused by the gradual accumulation of evidence and behavior was instead mathematically normative, we would expect the rating made on the second event of a gradually-accumulated-evidence trial, when the complete contents of the lakes are visible, to be identical to the rating made on the corresponding instantaneous evidence control trial. In other words, a difference between these ratings would be evidence of a bias.

Data from the instantaneous-evidence control trials for the Disconfirmatory condition were excluded from the analysis. This was done because it was the only condition in which the correct choice of focal hypothesis was inconsistent between the instantaneous-evidence trials and the gradual-evidence trials. Thus, it was not possible to produce an instantaneous-evidence control condition that was equivalent to the gradual-evidence condition in terms of support for the focal hypothesis and for each of its alternatives.

Results and discussion

All of the participants run were more than 80% accurate in selecting the most probable hypothesis to be the focal hypothesis. None of them were excluded from our analysis. We excluded any trials in which the focal hypothesis was less probable than one of its alternatives, whether it was selected by the participant or by the computer. We excluded an average of 3.8% (*SD* = 1.9%) of trials with self-selected focal hypotheses. The percentages were the same for trials with externally selected focal hypotheses, as

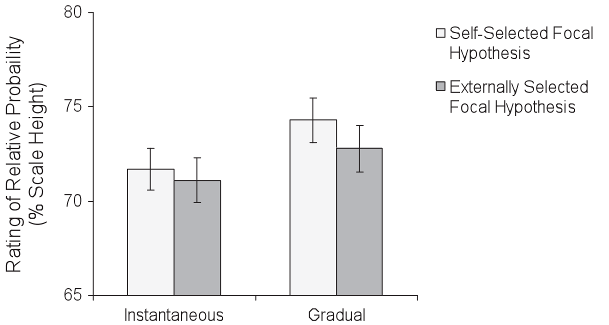

Fig. 3. Experiment 4: ratings of the relative probability that the focal hypothesis, rather than either of its alternatives, is true. As in Experiments 1–3, the focal hypothesis was judged to be more probable if it was self-selected than if it was externally selected. It was also judged to be more probable if the evidence was gradually accumulated than if it was all presented instantaneously.

accuracy was matched between the self and external conditions. We performed a 2 × 2 ANOVA to check for an interaction of Gradual Evidence Accumulation (whether evidence was gradually accumulated vs. all presented instantaneously) with Selection Type (self-selected vs. externally selected focal hypothesis). For gradually-accumulated-evidence trials, only the rating made at the end of the trial, once all of the evidence had become visible, was included in this analysis. This was compared to the one rating made on the corresponding instantaneous-evidence trial. As in Experiments 1–3, all statistical tests were repeated on rank transformed scores. All effects reported below to be significant in the original data were also significant in the rank transformed data. The data averaged across the Confirmatory and Neutral conditions are portrayed in Fig. 3. They are reported separately for each type of evidence in Table 5.

The ANOVA revealed a significant main effect of Gradual Evidence Accumulation, $F(1,41) = 27.95$, $p < .001$, $\eta^2 = .41$, involving a tendency to rate the focal hypothesis as more probable if evidence was gradually accumulated ($M = 37.7$), than if it was instantaneously presented ($M = 34.2$). This is consistent with the findings of previous research (Whitman & Woodward, 2011), in which evidence supporting a given focal hypothesis was seen to support it more strongly if gradually accumulated than if all presented instantaneously. There was also a significant self-selection bias in favor of self-selected focal hypotheses, $F(1,41) = 5.95$, $p < .05$, $\eta^2 = .13$, as in Experiments 1–3, with higher ratings for self-selected focal hypotheses ($M = 36.8$) than for externally selected ones ($M = 35.1$). The ratings of self-selected focal hypotheses were higher for 29 of the 42 participants. However, the interaction of Selection Type with Gradual Evidence Accumulation was not significant, $F(1,41) = 2.69$, $p = .11$. Finally, we found that self-selection bias and the effect of Gradual Evidence Accumulation were not correlated across participants, $r(42) = 0.12$, $p = .42$. These results do not support the notion that the two effects share a common underlying mechanism, such as repeated activation of mental representations.

General discussion

In four experiments, we found evidence for a self-selection bias whereby self-selected focal hypotheses were rated as being more probable than externally selected ones. This occurred despite the fact that: (1) the two types of focal hypothesis were matched in terms of mathematically normative probabilities, and (2) the overall accuracy of the computer in selecting the most probable hypothesis was matched, within-subjects, to the participant's accuracy in doing so. These results suggest that the cognitive operations involved in selecting a hypothesis lead to assignment of higher probability to that hypothesis.

Our secondary goal was to explore possible mechanisms by which the cognitive processes involved in selecting a focal hypothesis might lead to higher probability ratings. One of these was that the self-selection bias occurred because extra cognitive effort involved in focal hypothesis selection caused its mental representation to be more salient as a result of increased processing. We examined this possibility in Experiment 3 by manipulating the difficulty of identifying the most probable hypothesis. The results showed that the amount of cognitive effort required to select the focal hypothesis had no significant effect on the magnitude of self-selection bias. Another possible mechanism was that the self-selection bias occurred because of repeated activation of mental representations, leading to stronger subjective probabilities (a potential mechanism for the effects of gradual evidence presentation which was discussed in Whitman & Woodward, 2011). In Experiment 4, we manipulated whether evidence was gradually accumulated, so that evidence supporting the focal hypothesis

Table 5
Experiment 4: ratings of the relative probability that the focal hypothesis, rather than its alternatives, is true, as a function of the strength of the evidence supporting the focal hypothesis and of that supporting each alternative.

	Strength of evidence supporting each hypothesis (% fish in that lake of relevant color)			Predicted mathematically normative rating	Rating		Self-selection bias (self-ext)
	Focal	Alt. 1	Alt. 2		Self	External	
Confirmatory evidence							
Last rating, gradual	90	15	10	78.3	83.15	80.71	2.44
Only rating, instantaneous	90	15	10	78.3	80.95	80.33	0.62
Disconfirmatory evidence							
Last rating, gradual	70	77.5	47.5	35.9	50.37	49.08	1.29
Only rating, instantaneous	n/a	n/a	n/a	35.9	n/a	n/a	n/a
Neutral up							
Last rating, gradual	90	45	22.5	57.1	71.40	70.62	0.78
Only rating, instantaneous	90	45	22.5	57.1	69.63	68.89	0.74
Neutral down							
Last rating, gradual	70	35	17.5	57.1	68.33	67.02	1.31
Only rating, instantaneous	70	35	17.5	57.1	64.51	64.11	0.40

Note: In the Disconfirmatory Evidence condition, the balance of the evidence at the end of the trial refuted the focal hypothesis, so that support for the focal hypothesis (70%) was lower than support for its strongest alternative (77.5%). For this condition, it was not possible to include an instantaneous-evidence control condition with equivalent levels of support in which the focal hypothesis was self-selected, because it would be an error for a participant to select the lake with 70% fish of the relevant color as the focal hypothesis.

was processed repeatedly, or was instantaneously presented. We found no correlation between the effect of gradual vs. instantaneous evidence and self-selection bias, and thus found no evidence of a common underlying mechanism.

The results present two replicable judgment biases. These are: (1) a tendency to judge a self-selected hypothesis as more probable than one selected by an external source (self-selection bias) and (2) a tendency to weight gradually accumulated evidence more strongly than instantaneously presented evidence. This second bias replicates the results of our previous studies (Whitman & Woodward, 2011). Although both appear to be examples of repeated processing effects, whereby repeated processing of information consistent with a given hypothesis causes it to seem more salient and plausible, Experiment 4 suggests that if this is so, the underlying cognitive operations do not overlap.

One alternative interpretation stems from reports that the perceived value of information can be distorted prior to a decision (Bond, Carlson, Meloy, Russo, & Tanner, 2007; Brownstein, 2003; Dekay, Stone, & Miller, 2011; Russo, Carlson, & Meloy, 2006; Russo, Carlson, Meloy, & Yong, 2008), apparently in order to maximize consistency with the initially preferred option (Russo et al., 2008; Simon & Holyoak, 2002; Simon, Krawczyk, Bleicher, & Holyoak, 2008; Simon, Krawczyk, & Holyoak, 2004; Simon, Snow, & Read, 2004). In terms of our paradigm, this means that the perceived strength of evidence for the preferred hypothesis may have increased prior to focal hypothesis selection. Another potential interpretation is that the self-selection bias is a manifestation of the self-affirmation effect studied in social psychology (Brownstein, 2003; Steele, 1988; Steele, Spencer, & Lynch, 1993), whereby dissonance leads individuals to rate an option as being more desirable after choosing it than beforehand. Of course, pre-decision and post-decision biases are not mutually exclusive. They might even result from a common underlying mechanism, with the drive to maximize consistency being an inherent property of the cognitive system (Russo et al., 2008; Simon & Holyoak, 2002; Simon, Krawczyk, et al., 2004; Simon, Snow, et al., 2004; Simon, et al., 2008).

Finally, an alternative to repeated processing and cognitive consistency explanations is that the type of repeated processing that is involved in self-selection is qualitatively different than that involved in straight-forward repetition. When a preferred hypothesis is self-selected, the choice is a self-generated cognitive event (even though the hypothesis is not self-generated). It is clear that self-generated cognitive events are tagged with cognitive qualities that distinguish them from other-generated events (Johnson, Hashtroudi, & Lindsay, 1993). These qualities may lead to assignment of higher probability ratings.

Regardless of which of the above possible interpretations might contribute to selection bias (which will be an interesting direction for future research), the main contribution of the current work is that we demonstrated a selection effect in a paradigm with objectively quantifiable evidence. This was not the case with previous studies of self-selected material (Glockner, Betsch, & Schindler, 2010; Ronis & Yates, 1987; Sieck et al., 2007; Sniezek et al., 1990), consumer choice (Glockner & Betsch, 2008; Glockner et al., 2010; Russo et al., 2006, 2008) and legal decision-making (Simon & Holyoak, 2002; Simon, Krawczyk, et al., 2004; Simon, Snow, et al., 2004; Simon, et al., 2008). One other recent set of studies has investigated the effect of choice in a paradigm with objectively quantifiable evidence. However, that set of studies involved monetary gambles (Dekay et al., 2011). The current set of studies demonstrates that selection bias is not the result of confounds with individuals differences in personal experience, motivational leanings towards reward seeking and risk avoidance, or ease of memory retrieval. Rather, we demonstrate that some fundamental cognitive mechanism inherent to selecting a preferred focal hypothesis increases its perceived probability.

The direction of our self-selection bias is opposite to some previous experiments (Ronis & Yates, 1987; Sieck et al., 2007, 1990), which reported higher confidence in the correctness of other-selected material. However, as mentioned above as the motivation for this study, a number of variables confound the comparison of self-to-other selected trials, such as familiarity with the subject matter, fluency in memory retrieval, or completeness of memory retrieval. Moreover, the ratings made in previous studies were of confidence that the chosen answer had been correct, not probability comparisons as were carried out in the present study. These two processes (i.e., post hoc confidence ratings vs. probability comparisons) would seem to involve different cognitive operations, possibly contributing to the different self-selection effects. Finally, even when comparing our results to those of studies involving probability estimates rather than confidence estimates, we must consider that making a rating by moving a slider on a scale with the labels "definitely true" and "impossible" is not necessarily equivalent to making a numerical estimate of probability.

The findings of selection bias and of overweighting of gradually accumulated evidence in a paradigm requiring no long-term memory retrieval indicates the need to expand existing models of how choice affects confidence. These models describe how confidence ratings depend on fluency of retrieval from semantic and episodic memory (Ratcliff & Starns, 2009; Sieck, 2003; Sieck & Yates, 2001; Thomas, Dougherty, Sprenger, & Harbison, 2008). While there is no retrieval from long-term memory in our paradigm, there seem to be differences in the fluency of visual scene analysis between the self-selected and externally selected conditions. During the rating stage, the visual scene is analyzed more quickly in the self-selected condition than in the externally selected condition. This is presumably because some visual scene analysis was performed during the selection stage. The scene analysis performed during the rating stage is then faster/more fluent because it can build on information left in visual short-term memory from the previous scene analysis during the selection stage. Consequently, it may be that the selection bias is the result of more fluent visual scene analysis in the rating stage. An interesting direction for future research would involve adjustment of existing models to account for how the fluency of visual scene analysis is affected by choosing from one of multiple options (as opposed to the yes/no choices made in more basic, speeded perceptual discrimination tasks (Pleskac & Busemeyer, 2010).

A limitation of this study is that, although exclusion of semantic information and memory retrieval processes from our paradigm had the advantage of allowing us to reduce the cognitive confounds when comparing self- to other-selected trials, it may limit the generalizeability of our findings, as semantic information and memory retrieval processes are typically involved in opinion formation. Even when interpreting unfamiliar situations, individuals often make metaphorical comparisons to previous experiences or to semantic knowledge. For the sake of real-world validity, it will be necessary for future research to explore how the biases reported here and the memory-retrieval effects found in previous research, either compete or interact with each other in everyday belief/opinion formation. Similarly, while the removal of motivational factors involving risks and rewards allows us establish that selection bias can occur in the absence of those factors, this limits generalizability to real-world situations involving risky decision-making. Future research will be required to establish how selection bias is affected by such motivational factors.

The current studies suggest that the cognitive operations involved in self-selecting a hypothesis lead to assignment of higher probability to that hypothesis, and that this effect is independent of hypothesis selection difficulty and the rate of evidence accumulation. In the case of belief formation, self-selection bias could lead to a tendency to overvalue evidence confirming self-selected hypotheses. This could be a possible mechanism contributing to the holding of beliefs with weak supporting evidence, such as in horoscopes. This implies that difficulty accepting evidence that disconfirms beliefs may be partly due to the nature of the self-selection processes, which plays a role in elevating a hypothesis to belief status.

Acknowledgments

The authors thank Patrick Carolan and Jessica Ferguson-King for assistance with data collection and data management.

References

Beach, L. R. (1968). Probability magnitudes and conservative revision of subjective probabilities. *Journal of Experimental Psychology, 77*(1), 57–63.

Bond, S. D., Carlson, K. A., Meloy, M. G., Russo, J. E., & Tanner, R. J. (2007). Information distortion in the evaluation of a single option. *Organizational Behavior and Human Decision Processes, 102*(2), 240–254.

Brownstein, A. L. (2003). Biased predecision processing. *Psychological Bulletin, 129*(4), 545–568.

Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *American Statistician, 35*(3), 124–129.

Dekay, M. L., Stone, E. R., & Miller, S. A. (2011). Leader-driven distortion of probability and payoff information affects choices between risky prospects. *Journal of Behavioral Decision Making, 24*(4), 394–411.

Freeman, D., Pugh, K., & Garety, P. (2008). Jumping to conclusions and paranoid ideation in the general population. *Schizophrenia Research, 102*(1–3), 254–260.

Glockner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology – Learning Memory and Cognition, 34*(5), 1055–1075.

Glockner, A., Betsch, T., & Schindler, N. (2010). Coherence shifts in probabilistic inference tasks. *Journal of Behavioral Decision Making, 23*(5), 439–462.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3–28.

Koehler, D. J. (1994). Hypothesis generation and confidence in judgment. *Journal of Experimental Psychology: Learning Memory and Cognition, 20*(2), 461–469.

Moritz, S., Woodward, T. S., & Lambert, M. (2007). Under what circumstances to patients with schizophrenia jump to conclusions? A liberal acceptance account. *British Journal of Clinical Psychology, 46*, 127–137.

Peterson, C. R., Schneider, R. J., & Miller, A. J. (1965). Sample size and the revision of subjective probabilities. *Journal of Experimental Psychology, 69*(5), 522–527.

Pinel, P., Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage, 14*(5), 1013–1026.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review, 117*(3), 864–901.

Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review, 116*(1), 59–83.

Ronis, D. L., & Yates, J. F. (1987). Components of probability judgment accuracy – Individual consistency and effects of subject-matter and assessment method. *Organizational Behavior and Human Decision Processes, 40*(2), 193–218.

Russo, J. E., Carlson, K. A., & Meloy, M. G. (2006). Choosing an inferior alternative. *Psychological Science, 17*(10), 899–904.

Russo, J. E., Carlson, K. A., Meloy, M. G., & Yong, K. (2008). The goal of consistency as a cause of information distortion. *Journal of Experimental Psychology-General, 137*(3), 456–470.

Sieck, W. R. (2003). Effects of choice and relative frequency elicitation on overconfidence: Further tests of an exemplar–retrieval model. *Journal of Behavioral Decision Making, 16*(2), 127–145.

Sieck, W. R., Merkle, E. C., & Van Zandt, T. (2007). Option fixation: A cognitive contributor to overconfidence. *Organizational Behavior and Human Decision Processes, 103*(1), 68–83.

Sieck, W. R., & Yates, J. F. (2001). Overconfidence effects in category learning: A comparison of connectionist and exemplar memory models. *Journal of Experimental Psychology: Learning Memory and Cognition, 27*(4), 1003–1021.

Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and Social Psychology Review, 6*(4), 283–294.

Simon, D., Krawczyk, D. C., Bleicher, A., & Holyoak, K. J. (2008). The transience of constructed preferences. *Journal of Behavioral Decision Making, 21*(1), 1–14.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004a). Construction of preferences by constraint satisfaction. *Psychological Science, 15*(5), 331–336.

Simon, D., Snow, C. J., & Read, S. J. (2004b). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology, 86*(6), 814–837.

Sniezek, J. A., Paese, P. W., & Switzer, F. S. (1990). The effect of choosing on confidence in choice. *Organizational Behavior and Human Decision Processes, 46*(2), 264–282.

Speechley, W. J., Whitman, J. C., & Woodward, T. S. (2010). The contribution of hypersalience to the "jumping to conclusions" bias associated with delusions in schizophrenia. *Journal of Psychiatry & Neuroscience, 35*(1), 7–17.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology, 21*, 261–302.

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance – The role of affirmational resources. *Journal of Personality and Social Psychology, 64*(6), 885–896.

Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review, 115*(1), 155–185.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty – heuristics and biases. *Science, 185*(4157), 1124–1131.

Whitman, J. C., & Woodward, T. S. (2011). Evidence affects hypothesis judgments more if accumulated gradually than if presented instantaneously. *Psychonomic Bulletin & Review, 18*(6), 1156–1165.