Judgment and Decision Making, Vol. 17, No. 6, November 2022, pp. 1392–1421

Hypothesized drivers of the bias blind spot—cognitive sophistication, introspection bias, and conversational processes

David R. Mandel* Robert N. Collins† Alexander C. Walker‡
Jonathan A. Fugelsang§ Evan F. Risko¶

Abstract

Individuals often assess themselves as being less susceptible to common biases compared to others. This *bias blind spot* (BBS) is thought to represent a metacognitive error. In this research, we tested three explanations for the effect: The cognitive sophistication hypothesis posits that individuals who display the BBS more strongly are actually less biased than others. The introspection bias hypothesis posits that the BBS occurs because people rely on introspection more when assessing themselves compared to others. The conversational processes hypothesis posits that the effect is largely a consequence of the pragmatic aspects of the experimental situation rather than true metacognitive error. In two experiments (N = 1057) examining 18 social/motivational and cognitive biases, there was strong evidence of the BBS. Among the three hypotheses examined, the conversational processes hypothesis attracted the greatest support, thus raising questions about the extent to which the BBS is a metacognitive effect.

Keywords: bias blind spot, metacognition, cognitive ability, introspection, conversational processes

*Defence Research and Development Canada. Email: drmandel66@gmail.com. https://orcid.org/0000-0003-1036-2286.
†Defence Research and Development Canada. https://orcid.org/0000-0002-1714-7215.
‡HumanSystems Inc. https://orcid.org/0000-0003-1431-6770.
§HumanSystems Inc. https://orcid.org/0000-0002-6342-7023.
¶HumanSystems Inc. https://orcid.org/0000-0001-5702-0350.

This research was supported by Canadian Safety and Security Program project CSSP-2018-TI-2394. We thank Brooke MacLeod for her research assistance. Experiments were not preregistered. Data and supplementary materials are available via the Open Science Foundation at https://osf.io/dzkxy/.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

1 Introduction

Judgment and decision-making can be systematically biased in several ways (Gilovich et al., 2002; Kahneman et al., 1982). Bias can lead to poor judgments and decisions in domains such as intelligence analysis (Dhami et al., 2019; Morewedge et al., 2015), medicine (Arkes, 2013; Bornstein & Emler, 2001) and forensics (Kassin et al., 2013) to name a few. This is compounded by the fact that individuals, including children (Elashi & Mills, 2015; Hagá et al., 2018), often report that they are less susceptible to biases than others — a phenomenon known as the *bias blind spot* (BBS; Pronin et al., 2002; Pronin & Kugler, 2007; Scopelliti et al., 2015; West et al., 2012).

People exhibit a BBS in judgments about their own and others' abilities to objectively process information about criminal cases (Jones et al., 2018). As well, forensic experts whose opinions play a significant role in criminal proceedings perceive bias as a problem in other forensic experts, while identifying much less bias in their own judgments (e.g., Kukucka et al., 2017; Zappala et al., 2018). Given the important role that awareness of our decision-making limitations likely has in correcting them, having a better understanding of the nature of this metacognitive bias is critical. However, the cause of the BBS remains unclear. This article addresses that question while also introducing various methodological refinements to the research area.

Pronin et al. (2002) provided the first examination of the BBS. In that work, participants read short descriptions of various biases that were largely social/motivational in nature (e.g., self-serving bias and the halo effect). Participants were asked to rate their own susceptibility to each presented bias or the susceptibility of the average American, classmate, or fellow travellers in an airport (across separate studies). Individuals reported less susceptibility to the biases examined when judging themselves rather than others. The BBS was subsequently demonstrated by Ehrlinger et al. (2005) and effects from Pronin et al. (2002) were successfully replicated in a series of preregistered experiments by Chandrashekar et al. (2021).

West et al. (2012) extended earlier social psychological work on the BBS to the cognitive domain using classic examples from the heuristics and biases literature (e.g., outcome bias and base-rate neglect). Again, the authors asked participants to read a brief description of a bias and then rate their susceptibility to the bias and that of an average person (or average student from the same university). West et al. (2012) found evidence of a BBS, which was observed across all bias items, and the individual bias blind spots were inter-correlated.

Lastly, in order to develop and validate a BBS scale, Scopelliti et al. (2015) examined the BBS across a number of different biases, most of which were social/motivational in nature, similar to those studied by Pronin et al. (2002). Scopelliti et al. (2015) found a significant BBS for each bias item in the scale and, once again, individual bias blind spots were inter-correlated. Thus, the BBS appears to be a robust and replicable phenomenon that generalizes across several different social/motivational and cognitive biases.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

1.1 Potential explanations

Although the BBS is robust, its cause remains unclear. In particular, individual differences in the expression of the BBS still require explanation. In the present article, we consider three possible explanations, which we view as viable contenders. We do not, however, view them as mutually exclusive, nor do we assume that the BBS is a monocausal phenomenon. To the contrary, we suspect that this phenomenon may be influenced by multiple factors, a subset of which we explore in the present research.

1.1.1 The cognitive sophistication hypothesis

One hypothesis we explore is that variation in the BBS reflects corresponding differences in actual cognitive performance, namely, people who show stronger bias blind spots may actually be *less* biased than others. This cognitive sophistication hypothesis receives some support from West et al. (2012), who examined the relation between the BBS and measures of cognitive sophistication including self-reported scholastic aptitude test (SAT) scores, the Cognitive Reflection Test (CRT; Frederick, 2005), Need for Cognition (NFC; Cacioppo et al., 1996), and the Actively Open-minded Thinking (AOT) scale (Stanovich & West, 1997, 2007). Interestingly, the various measures of cognitive sophistication were all positively correlated with the BBS. That is, supporting the cognitive sophistication hypothesis, individuals who scored higher on measures of cognitive sophistication had *larger* bias blind spots. However, contrary to the cognitive sophistication hypothesis, West et al. (2012) found that performance on actual measures of cognitive bias were generally unrelated to cognitive sophistication, an overall measure of their BBS, or to the specific BBS measures associated with the actual biases (e.g., the relation between the anchoring bias blind spot and the actual anchoring bias). The positive relation between the BBS and cognitive sophistication also has not been consistently observed. For instance, Scopelliti et al. (2015) found no relation between scores on their BBS scale and self-reported verbal SAT scores and a small negative correlation with self-reported math SAT scores. However, since West et al. (2012) used only cognitive bias items and Scopelliti et al. (2015) used only social/motivational items, it is possible that the discrepancy is attributable to the different domains. Another limitation of all previous studies examining the link between the BBS and cognitive sophistication is that the primary measure — cognitive ability — was always based on self-reported SAT scores. The reporting of SAT scores could be not only error prone, but also systematically biased in ways that relate to the BBS.

In the present research, we addressed the methodological issues that have cast ambiguity on the cognitive sophistication hypothesis by administering both types of bias items (i.e., cognitive and social) while also collecting data on cognitive sophistication and performance on actual bias tasks corresponding to our specific BBS items (Experiment 2). A novel feature of our research (in both Experiments 1 and 2) is that instead of collecting self-reported SAT scores as a measure of cognitive ability, we use a well-validated measure of cognitive ability,

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

the International Cognitive Ability Resource (ICAR; Condon & Revelle, 2014; Dworak et al., 2021). Additionally, we collected AOT scores in both experiments, which enabled an assessment of whether this cognitive style measure adds unique variance to the prediction of the BBS. AOT assesses an individual's willingness to evaluate evidence contrary to their beliefs as well as openness to alternative perspectives (Baron et al., 2015). AOT is positively related to accuracy in a variety of judgment tasks (Haran et al., 2013; Mandel et al., 2020) and negatively related to some cognitive biases (Toplak et al., 2017).[1] Finally, in addition to examining the relation between the BBS and these measures of cognitive ability, we also examine the relation between the individual self and other ratings, which has not been explored in earlier studies. Since the BBS is a difference of related self and other assessments, this more detailed correlational analysis is vital to unpacking the basis of the previously reported correlations between cognitive sophistication and the BBS.

1.1.2 The introspection bias hypothesis

An alternative explanation for the BBS focuses on the metacognitive strategies that participants use when making bias susceptibility judgments (Ehrlinger et al., 2005; Pronin et al., 2004; Pronin & Kugler, 2007). In one experiment, Pronin and Kugler (2007) presented participants with descriptions of social/motivational biases and asked them to judge themselves or others' susceptibility to bias. In addition, participants indicated the extent to which they used two different strategies for assessing bias: (1) "trying to [get inside my head/get inside the heads of particular Harvard students] to find evidence of the sorts of thoughts and motives that could underlie this tendency" (p. 3; Pronin & Kugler, 2007) and (2) "considering how well this description fits the way that people in general tend to behave" (p. 3; Pronin & Kugler, 2007). Results showed an interaction effect such that when participants judged themselves, they relied more on introspection, but when participants judged others, they relied more on considering how well the bias described people in general. Pronin and Kugler (2007) interpreted this pattern as providing support for an *introspection bias hypothesis* in which individuals underestimate their bias susceptibility because the biases typically examined in the literature are not easily detectable by introspection. Conversely, these biases may be more readily detected when examining (observable) behavior — a strategy that is more commonly used (or at least reported to be used) when assessing others' susceptibility to bias.

While the evidence in Pronin and Kugler (2007) and elsewhere (Ehrlinger et al., 2005; Pronin et al., 2004, 2007) supports the introspection bias hypothesis, to the best of our knowledge, this result has not yet been replicated. Moreover, the biases used in that work were exclusively within the social/motivational realm. Accordingly, we attempted to replicate the original result and extend it using both social/motivational and cognitive bias items (Experiment 2). We also used a novel procedure that better segregated participants' self

[1] In the interest of keeping the survey completion time lower, we did not collected data on CRT or NFC, but these measures tend to be positive correlated with AOT and cognitive ability.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

and other judgments. In our spaced randomized-block elicitation procedure, participants provide self and other ratings in separate blocks whose order was randomized, and which are spaced by intervening tasks. This combines the statistical power of repeated-measures designs with a closer approximation of separate evaluability for self- and other-bias ratings. Moreover, because we collect participants' responses in the same blocks, our design affords greater separability of these responses for the self and others than had been possible in earlier within-subject tests of the introspection bias hypothesis.

1.1.3 The conversational processes hypothesis

A third explanation we examine is that the BBS is not principally a metacognitive effect, but rather the result of conversationally pragmatic processes that occur in the context of typical BBS tasks — namely, an experimental artefact. That is, the BBS might reflect a conversational process in which the information that participants are given prior to providing their ratings are assumed by participants to be true and relevant to the task (Hilton, 1995; Sperber & Wilson, 1986). In BBS studies, participants are explicitly informed about the behavior of people, in general, and then, when assessing others, they are asked to consider the *average* respondent. Given that they are told that the average individual is biased, they may justifiably come to believe that the average survey respondent is biased. In contrast, when participants provide self-assessments, there is no logical coupling between their assessment and the earlier claim that most people may be biased – indeed, this statement does not imply that *I* am biased.

If the *conversational processes hypothesis* is correct, then we would expect to observe the following: First, we would expect to find greater variability among self-assessments than among other-assessments of bias. This is because other-assessments should be more tightly constrained by the information that is provided than self-assessments. Second, given the content of that constraining information (i.e., people are biased), we would expect to find that a greater proportion of participants's other-assessments fall above the midpoint of the rating scale, because this pattern of response represents agreement with what participants had been told was true, in general. We test these predictions in Experiments 1 and 2.

2 Experiment 1

Participants completed the 14 BBS items used in Scopelliti et al. (2015). In addition, participants completed two measures of cognitive sophistication (i.e., ICAR and AOT). As noted earlier, whereas West et al. (2012) found a positive correlation between the BBS and a self-reported measure of cognitive ability, Scopelliti et al. (2015) did not. West et al. (2012) also found a positive relation between the BBS and AOT. Experiment 1 allowed us to attempt to conceptually replicate these relations across a new set of biases (i.e., those from Scopelliti et al., 2015) and using a different (and much shorter) measure of AOT than West

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

et al. (2012) used. Additionally, Experiment 1 permitted an initial test of the conversational processes hypothesis.

2.1 Method

2.1.1 Participants

Participants completed Experiment 1 online via Qualtrics Panels. This study was open to Canadian and U.S. citizens between the ages of 18 and 60 years of age who self-reported English as their first language. Only participants who passed a pre-survey attention check and took at least 500 seconds to complete the experiment were included in the sample (N = 767). The attention check item was as follows: "The survey that you are about to enter is longer than average, and will take about 30 to 60 minutes. There is a right or wrong answer to some of the questions, and your undivided attention will be required. If you are interested in taking this survey and providing your best answers, please select "No" below. Thank you for your time!" Participants who did not respond "No" were excluded. Additionally, 3 participants were excluded for missing data and 111 were excluded for failing the following in-survey attention check: "In the following alphanumeric series, which letter comes next? A B C D E" with options (1) *I* (2) *H* (3) *G* (4) *F* (5) *J*.[2]

Thus, the final sample included 653 participants (57% female; M_{age} = 39.92[3], SD_{age} = 11.96; 66% obtained a college diploma or higher; 311 Canadian citizens, 321 US, and 21 dual). Sensitivity power analyses conducted using G*Power (Faul et al., 2007) indicated that our retained sample yielded 80% power to detect small effects (d = 0.11 and r = .11) for planned analyses.

Following completion of Experiment 1, participants received compensation from the Qualtrics panel provider.

2.1.2 Materials and procedure

Participants first completed six demographic questions (i.e., age, sex, education level, nationality, citizenship, and first language). They then completed a series of tasks in a counterbalanced order. The tasks included the BBS items, ICAR, AOT, and an estimation task that required providing estimates in response to a series of questions on various topics (e.g., general knowledge) and constructing intervals around those estimates (see Mandel et al., 2020).[4] The median completion time of the experiment was 1,489 seconds (approximately 25 minutes), and participants were debriefed at the end.

[2]Missing responses on ICAR were treated as incorrect answers rather than as missing data. An additional item ("The value of a quarter is what percentage of a dollar?") was administered but not used for attention screening as it was later deemed by the authors to be closer to a numeracy test item.

[3]Participants that did not provide an age in years (e.g., instead provided a year of birth) were not included in this calculation.

[4]The present report focuses on responses to the BBS items, ICAR, and AOT, which were not the focus of the previous report by Mandel et al. (2020).

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Bias blind spot We divided Scopelliti et al.'s (2015) items into two sets of seven, which corresponded to the first and second halves of Scopelliti et al.'s scale (Appendix A). Set 1 comprised the action-inaction bias, the bandwagon effect, the confirmation bias, the disconfirmation bias, diffusion of responsibility, escalation of commitment, and fundamental attribution error. Set 2 comprised the halo effect, ingroup favoritism, the ostrich effect, projection bias, self-interest bias, self-serving bias, and stereotyping. Each participant completed either Set 1 or Set 2. Participants were given descriptions of the various biases (e.g., confirmation bias), and were asked: (1) "To what extent do you believe that *you* show this effect or tendency?" and (2) "To what extent do you believe the *average survey respondent* shows this effect or tendency?" Responses were provided on a 1 (*not at all*) to 7 (*very much so*) rating scale. Following Scopelliti et al. (2015), we did not counterbalance or randomize the order in which self and other ratings were provided, with self-ratings being first. The mean difference between the self and other ratings (which ranged from −6 to 6) was referred to as the BBS score and served as the dependent variable. As in Pronin et al. (2002) and Scopelliti et al. (2015), participants were also asked a final general question about their belief that objective measures would find them to be more, less or as biased as they rated themselves. Specifically, they read the following:

> Studies have shown that on the whole, people show a "better than average" effect when assessing themselves relative to other members within their group. That is, 70–80% of individuals consistently rate themselves "better than average" on qualities that they perceive as positive, and conversely, evaluate themselves as having "less than average" amounts of characteristics they believe are negative. For the purposes of our study, it would be useful to know the accuracy of your self-assessments on the previous pages. Please indicate how you would be rated on the relevant dimensions by the "most accurate, valid, and objective" resources available.

Following this, they were provided with three response options:

1. The objective measures would rate me lower on positive characteristics and higher on negative characteristics than I rated myself.

2. The objective measures would rate me neither more positively nor more negatively than I rated myself.

3. The objective measures would rate me higher on positive characteristics and lower on negative characteristics than I rated myself.

These responses were coded from 1 to 3, respectively.

Cognitive ability Participants completed an 8-item measure of cognitive ability using items selected from the ICAR (Condon & Revelle, 2014; The International Cognitive Ability Resource Team; https://icar-project.com/). Two items were selected from the Verbal Reasoning (VR.4 and VR.19), Letter and Number Series (LN.7 and LN.33), Matrix Reasoning (MR.45 and MR.47), and Three-Dimension Rotation (R3D.4 and R3D.6) sections.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Participants were provided with a problem description and six answers from which to select the correct response. Response options allowing participants to state that 'no option was correct' or that they 'did not know the answer' were removed. Participants' average ICAR scores (i.e., the proportion of correct responses) were used in subsequent analyses.

Actively open-minded thinking Participants completed an 8-item AOT measure (Baron et al., 2015). Participants rated their agreement or disagreement on a scale from −2 to 2 with statements about evaluating evidence (e.g., "One should disregard evidence that conflicts with one's established beliefs"). Participants' average AOT scores were used in subsequent analyses.

2.2 Results and discussion

2.2.1 Bias blind spot

We first ascertained the unidimensionality and reliability of an averaged BBS score based on the two sets of items used. We submitted each set of BBS items to a one-factor confirmatory factor analyses (CFA) with BBS as the sole predicted factor and using the *lavaan* package in the *rStudio* programming environment running *r* code. We used a robust maximum likelihood (ML) estimator with a nonlinear minimization subject to box constraints (NLMINB) optimization method. Our criteria for ideal model fit were a non-significant χ^2 test of fit ($p > .05$), a comparative fit index (CFI) $\geq .90$, a Tucker Lewis index (TLI) ≥ 0.95, a root mean square error of approximation (RMSEA) ≤ 0.08, and a standardized root mean square residual (SRMR) ≤ 0.08 (Kline, 2005; Hooper et al., 2008). The reliability of the response scales was ascertained if $\omega_t \geq .70$ (McDonald, 1999; Revelle & Condon, 2019). Both sets of items satisfied each of our criteria for a reliable fit to a unidimensional structure (see Table 1). Accordingly, we averaged BBS scores as the primary dependent variable for further analysis.[5]

Table 1: Results of CFA for Experiment 1.

	Goodness of Fit Test			Fit Indices				Reliability
Model	χ^2	*df*	*p*	CFI	TLI	RMSEA	SRMR	ω_t
BBS Set 1	12.71	14	.550	1.000	1.005	.000	.028	.79
BBS Set 2	17.77	14	.218	.980	.970	.038	.041	.73

Note. $N = 332$ for BBS Set 1, $N = 321$ for BBS Set 2.

[5]Figures S1 and S2 in the supplementary materials provide additional information on the relations among individual BBS items in Sets 1 and 2, respectively. Figures S1 and S2 also show the distribution of BBS for each item in Sets 1 and 2.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Next, we examined whether participants demonstrated the BBS. As Table 2 shows, replicating Scopelliti et al. (2015), participants demonstrated the BBS across all items in both sets (all *p*s < .023). Overall, the BBS effect size was medium, and ranged from the smallest magnitude for the halo effect ($d = 0.12$) to the largest magnitude for the self-serving bias ($d = 0.46$).

Table 2: Bias blind spot ratings in Experiment 1.

	Other		Self						
Bias Type	*M*	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*	95% CI
Average	4.73	.97	4.20	1.27	13.85	652	<.001	0.45	0.39, 0.52
Set 1									
Action-inact.	4.61	1.37	4.34	1.74	3.37	342	.001	0.18	0.07, 0.28
Bandwagon	4.80	1.34	4.08	1.88	7.80	342	<.001	0.44	0.32, 0.55
Confirm.	4.80	1.39	4.16	1.73	7.53	342	<.001	0.42	0.31, 0.54
Disconfirm.	4.81	1.34	4.47	1.62	4.72	342	<.001	0.23	0.14, 0.33
Diff. Resp.	4.89	1.39	4.20	1.78	7.68	342	<.001	0.44	0.33, 0.56
Esc. Comm.	4.80	1.34	4.06	1.85	9.26	342	<.001	0.46	0.35, 0.56
Attr. Error	4.59	1.37	3.80	1.74	9.82	342	<.001	0.50	0.39, 0.61
Set 2									
Halo	4.78	1.28	4.60	1.50	2.30	331	.022	0.12	0.01, 0.24
Ingroup Fav.	4.48	1.44	3.98	1.73	6.50	331	<.001	0.31	0.21, 0.41
Ostrich	4.61	1.40	4.37	1.74	3.26	331	.001	0.16	0.07, 0.25
Projection	4.73	1.27	4.44	1.63	3.35	331	.001	0.18	0.06, 0.29
Self-Interest	5.08	1.34	4.66	1.52	5.55	331	<.001	0.29	0.18, 0.40
Self-Serving	4.78	1.29	4.05	1.68	9.38	331	<.001	0.46	0.36, 0.57
Stereotyping	4.60	1.46	3.83	1.96	8.31	331	<.001	0.43	0.32, 0.54

Note. Tests are based on paired-samples *t*-tests comparing self and other ratings.

Providing initial support for the conversational processes hypothesis, self-assessments were more variable than other-assessments for each bias examined and, on average, the heterogeneity in variance was significant (Levene's $F = 44.66$, $p < .001$). In addition, as predicted by the conversational processes hypothesis, a significant majority of participants (74.4%, 95% CI [70.9%, 77.6%]) had a mean other-bias rating above the scale midpoint. This percentage was significantly greater than the corresponding percentage for mean self-bias ratings (49.6%, 95% CI [45.8%, 53.4%]; McNemar $\chi^2 = 122.27$, $p < .001$).

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

On average, participants judged that their self-assessments were unbiased (M = 2.03, SD = 0.65). Fully, 58.3% assessed themselves to be unbiased. The remainder of the sample was near evenly split with 19.3% of participants having assessed their behavior to be worse than they rated it and 22.5% of participants having assessed their behavior to be better than they rated it. A one-tail chi-square test comparing the latter two proportions was not significant (p = .182).

Consistent with West et al. (2012), overall BBS was positively correlated with ICAR (r = .20, p < .001) and AOT (r = .16, p < .001). ICAR (M = 0.37, SD = 0.24) and AOT (M = 0.43, SD = 0.58) were also positively correlated (r = .37, p < .001). To assess whether ICAR and AOT contributed independently to predicting BBS, we regressed overall BBS on these measures. Both ICAR (B = 0.64, SE = 0.17, β = 0.16, t = 3.83, p < .001) and AOT (B = 0.17, SE = 0.07, β = 0.10, t = 2.46, p = .014) were significant predictors. To better understand the basis for the correlations between BBS and ICAR and AOT, we correlated the latter two measures with the average assessment of self- and other-bias. ICAR was correlated with self-bias ratings (r = –.17, p < .001), but not with other-bias ratings (r = –.02, p = .62). AOT was significantly correlated with self-bias ratings (r = –.22, p < .001) as well as with other-bias ratings (r = –.12, p = .002). Therefore, the correlations observed between BBS and cognitive sophistication measures appear to be largely driven by the self-bias ratings.

3 Experiment 2

Experiment 2 extends Experiment 1 in several ways. First, we again examined the relation between cognitive ability, AOT, and the BBS using a modified subset of biases and ICAR items. In Experiment 2 we examined the effect of bias type by comparing three distinct types of items. Two of the biases used in Experiment 2 (i.e., self-serving bias and the halo effect) were social/motivational in nature, as in Scopelliti et al. (2015), and four of the biases were cognitive biases taken from the heuristics and biases and reasoning literatures (i.e., belief bias, outcome bias, conjunction bias, and anchoring bias) as in (except for belief bias) West et al. (2012). Additionally, the four cognitive biases were selected because two of them (i.e., belief bias and outcome bias) have been shown to relate to cognitive ability and two of them (i.e., conjunction bias and anchoring bias) have been shown not to relate to cognitive ability (Stanovich & West, 2008). The use of bias items that either were or were not related to cognitive ability in earlier research provides a novel test of the basis for the relation between the BBS and cognitive ability. That is, if the relation between the BBS and cognitive ability reflected genuine insight into one's diminished susceptibility to the biases described, then the relation between the BBS and cognitive ability should be present only (or at least more strongly) in a set of biases that are related to cognitive ability. Finally, in Experiment 2 we separated blocks containing self and other bias judgments and manipulated the order of these blocks. In previous work, self and other judgments were either elicited in immediate succession (i.e., prompting a joint evaluation mode) or between

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

participants (i.e., prompting a single evaluation mode). Here, we assessed whether this novel design feature influences expression of the BBS.

We also examined the relation between the BBS and behavioral measures of actual biases in Experiment 2. In addition, we extend previous work by including a bias that has not previously been examined in this manner (i.e., belief bias in syllogistic reasoning) and by also including items that index susceptibility to social/motivational biases (i.e., self-serving bias and the halo effect). Thus, we can assess the relation between bias blind spots and the expression of these biases across a varied set of biases.

Lastly, in addition to asking participants to rate their susceptibility and others' susceptibility to biases, Experiment 2 had participants report on the strategy they used to make these assessments. As noted earlier, Pronin and Kugler (2007) proposed that a causal basis of the BBS is that participants use different strategies when rating themselves versus others. According to the introspection bias hypothesis, when participants rate themselves, they rely more on introspection, but when rating others they rely more on considering how accurately the bias described how people behave, in general. Here, we attempt to replicate and extend this finding to a new set of biases. Participants were asked about their use of each strategy for all six biases. Critically, the social/motivational biases used in Experiment 2 were the same as those used in Pronin and Kugler (2007), while the cognitive biases represent an opportunity to test the degree to which the introspection bias hypothesis extends to cognitive items.

3.1 Method

3.1.1 Participants

A sample of 427 participants completed Experiment 2 online via Qualtrics Panels. Consistent with Experiment 1, this study was open to Canadian and U.S. citizens between the ages of 18 and 60 years of age who self-reported English as their first language. Only participants who passed two pre-survey attention checks and took at least 500 seconds to complete the experiment were included in the sample. Of the pre-survey attention checks used in Experiment 2, one involved asking participants to select, from a list of movie titles, the one that would come 4th if they were arranged alphabetically (adapted from Littrell & Fugelsang, 2021). The other required the participant to type the statement, "bot a not am I" in reverse order in an open text box. If participants failed either of these items, Qualtrics Panels excluded them from the sample. Additionally, 7 participants were removed for providing incomplete responses and 16 were removed for failing an in-survey attention check adapted from Oppenheimer et al. (2009) and described below.[6] Thus, the final sample

[6]Our variation in the use of attention checks was guided by the belief that it would be better to screen on multiple items initially and use ones that involved quite basic tasks without lures. Conversely, we suspected that an item such as that based on Oppenheimer et al.'s instructional manipulation task would be better at assessing inattention to instructions at a stage where boredom may have set in. These variations were not based on any analysis of data exclusions in Experiment 1.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

included 404 participants (51% female; M_{age} = 44.45, SD_{age} = 10.52; 68.1% obtained a college diploma or higher; 209 Canadian citizens, 189 US, and 6 dual). Sensitivity power analyses indicated that our retained sample yielded 80% power to detect small effects (d = 0.14, r = .14) for planned analyses. As in Experiment 1, participants received compensation from the Qualtrics panel provider upon study completion.

3.1.2 Materials and procedure

As in Experiment 1, participants began Experiment 2 by responding to a set of demographic questions. Next, they completed six tasks designed to assess their biases. Participants were randomly assigned to one of two item groupings (Group 1 and 2). Group 1 completed Set A of the self-serving, conjunction, and outcome bias items and Set B of halo, anchoring, and belief bias items, with the reverse being true for those assigned to Group 2. Following the completion of all bias items, participants completed the ICAR along with the letter sequence attention check. Participants then completed the first of two BBS item blocks (self or other), followed by the AOT (identical to Experiment 1) and an "Instructional Manipulation Check" adapted from Oppenheimer et al. (2009) whereby under the cover of a question about sports participation, participants were instructed to ignore the main question and click a button to proceed to the next screen. After completing these intervening tasks, participants completed the second BBS item block. BBS items of the same type (i.e., social, cognitive biases related to cognitive ability, cognitive biases unrelated to cognitive ability) were always presented together, with the order of presentation by type randomized. The median completion time of the experiment was 1,368 seconds (approximately 23 minutes), and participants were debriefed at the end.

Bias blind spot We assessed the BBS using six items (Appendix B) grouped into three types of biases: (1) two social/motivational biases (self-serving bias and the halo effect), (2) two cognitive biases shown by Stanovich and West (2008) to be related to cognitive ability (outcome bias and belief bias), and (3) two cognitive biases shown by Stanovich and West (2008) to be unrelated to cognitive ability (anchoring bias and conjunction bias). All BBS items were presented within both self and other blocks in which participants were asked to make judgments with respect to themselves or others (i.e., the average survey respondent), respectively. The order of the self and other blocks was randomized across participants. For each BBS item, participants were given a brief description of the relevant bias and asked three questions: (1) "To what extent do you believe [that you show/the average survey respondent shows] this effect or tendency?" (2) How much did you try to 'get inside your head' ['get inside the head of the average survey respondent'] to find evidence of the sorts of thoughts and motives that could underlie this tendency? (3) How much did you try to consider how well this description fits the way that people in general tend to behave? Responses to these three questions were provided on a 7-point scale, ranging from 1 (*not at all*) to 7 (*very much*). Responses to the first of these questions was used

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

to assess the BBS, whereas the responses to the latter two questions were used to assess strategy use. At the end of each BBS item block, participants responded to an item asking them how well they understood the bias descriptions: "I understood the descriptions of the tendencies provided." Responses to this item were provided on a 6-point scale that ranged from *completely disagree* (subsequently dummy coded as 1) to *completely agree* (dummy coded as 6). We created total and subset bias scores by averaging the total bias difference score (self minus other) across all items.

Bias items We used six bias items corresponding to the six BBS items (i.e., self-serving bias, halo effect, outcome bias, belief bias, conjunction fallacy, and anchoring bias). Bias was assessed between-subjects, and as such, two sets of each bias item were created. Bias items were presented in a random order for each participant. Appendix C provides a full description of each task and the scoring procedure.

Cognitive ability Participants completed ten ICAR items (Condon & Revelle, 2014; The International Cognitive Ability Resource Team; https://icar-project.com/), four of which were administered in Experiment 1, plus six new items. Five items each were selected from the Verbal Reasoning and Matrix Reasoning items: VR.14, V4.04, VR.16, VR.17, VR.19, MR.45, MR.46, MR.43, MR.47, and MR.55. As in Experiment 1, each item presented participants with a problem description and six answers from which to select the correct response. Response options of 'None of these' and 'I don't know' were removed. These 10 items were also presented in a random order. As in Experiment 1, the dependent variable here was the proportion of correct responses.

3.2 Results and discussion

3.2.1 Bias blind spot

The BBS items in Experiment 2 were submitted to a CFA. We used the same methodology as in Experiment 1 with one important exception. In Experiment 2, we tested one-, two-, and three-factor models. Specifically, the two-factor model includes a BBS Social/Motivational factor (halo effect and self-serving bias) and a BBS Cognitive factor (outcome bias, belief bias, conjunction fallacy, and anchoring bias). The three-factor model includes the aforementioned BBS Social/Motivational factor, a BBS Cognitive Bias Related to Cognitive Ability (Cognitive-Related) factor (outcome bias and belief bias), and a BBS Cognitive Bias Unrelated to Cognitive Ability (Cognitive-Unrelated) factor (conjunction fallacy and anchoring bias).

As Table 3 shows, each of the proposed models provided a reliable fit for the data. However, we found no statistical justification for increasing model complexity beyond a simple one-factor BBS model. The two-factor model did not significantly improve model fit compared to the one-factor model. Likewise, the three-factor model did not

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

significantly improve model fit compared to the two-factor model (i.e., $p > .05$ for $\Delta\chi^2$ for increasingly complex models in both cases; see Table S1 in supplementary materials). However, giventhat all three models exhibit sound properties and there is an a priori theoretical interest in the finer distinctions captured in the two- and three-factor models, we computed BBS scales derived from each model by averaging the relevant subsets of items.

Table 3: Results of CFA for Experiment 2.

	Goodness of Fit Test			Fit Indices				Reliability
Model	χ^2	*df*	*p*	CFI	TLI	RMSEA	SRMR	ω_t
One factor	4.91	9	.842	1.000	1.031	.000	.021	.71
Two factor	4.87	8	.771	1.000	1.026	.000	.021	.94
Three factor	3.73	6	.714	1.000	1.027	.000	.019	.95

Note. $N = 404$.

Consistent with Experiment 1, participants believed that they were less biased than the average survey respondent. In fact, as Table 4 shows, bias blind spots were observed for each of the six biases investigated. Consistent with Pronin et al. (2002), the magnitude of participants' BBS did not differ as a function of the order of presentation of self and other items ($t[402] = 0.82$, $p = .41$). As in Experiment 1,the size of participants' bias blind spots was positively correlated across all six biases (see Figure S3 in the supplementary materials). A one-way repeated-measures analysis of variance (ANOVA) showed a significant effect of bias type on BBS scores ($F[2, 425] = 5.16$, $p = .006$, $\eta^2_p = .024$). Post-hoc tests revealed that the social/motivational BBS scores were significantly greater than BBS scores for the cognitive items that were previously shown to have biases that were unrelated to cognitive sophistication measures (i.e., Cognitive-Unrelated items). However, the Cognitive-Related BBS scores did not significantly differ from either of these other scores.

Consistent with Experiment 1, and providing additional support for the conversational processes hypothesis, the variance of self-bias assessments was greater than that of other-bias assessments for each of the six biases examined, and overall, the variance between self and other ratings was once again significantly heterogeneous (Levene's $F = 6.52$, $p = .011$). Moreover, as predicted by the conversational processes hypothesis, a significant majority of participants (77.7%, 95% CI [73.4%, 81.5%]) had a mean other-bias rating above the scale midpoint. This percentage was significantly greater than the corresponding percentage for mean self-bias ratings (61.9%, 95% CI [57.1%, 66.5%]; McNemar $\chi^2 = 38.16$, $p < .001$).

3.2.2 Bias blind spot, cognitive sophistication and bias

Consistent with the findings of Experiment 1, overall BBS scores were positively correlated with ICAR ($r = .20$, $p < .001$) and AOT ($r = .19$, $p < .001$). As in Experiment 1, these

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Table 4: Bias blind spot ratings in Experiment 2.

	Other		Self					
Bias Type	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*	95% CI
All Biases	4.84	1.13	4.36	1.28	8.92	<.001	0.40	0.31, 0.49
Social/Motivational	5.00	1.27	4.41	1.54	8.24	<.001	0.42	0.31, 0.52
Halo Effect	4.97	1.49	4.36	1.78	7.23	<.001	0.37	0.26, 0.47
Self-Serving Bias	5.03	1.45	4.46	1.68	6.60	<.001	0.37	0.25, 0.48
All Cognitive	4.76	1.20	4.33	1.31	7.64	<.001	0.34	0.25, 0.43
Cognitive-Related	4.91	1.33	4.45	1.48	7.64	<.001	0.32	0.23, 0.42
Outcome Bias	5.11	1.48	4.60	1.75	6.16	<.001	0.31	0.21, 0.41
Belief Bias	4.71	1.64	4.30	1.77	4.58	<.001	0.24	0.13, 0.34
Cognitive-Unrelated	4.61	1.33	4.21	1.46	5.91	<.001	0.28	0.18, 0.37
Conjunction Fallacy	4.75	1.48	4.50	1.67	3.18	.002	0.16	0.06, 0.27
Anchoring Bias	4.46	1.63	3.93	1.76	6.16	<.001	0.31	0.21, 0.41

Note. $df = 403$. All inferential statistics are based on paired-samples *t*-tests comparing ratings provided across self and other items.

correlations were invariably small. Given that ICAR ($M = 0.53$, $SD = 0.27$) and AOT ($M = 0.61$, $SD = 0.56$) were also correlated, we regressed overall BBS scores on these variables. As in Experiment 1, both ICAR scores ($B = 0.58$, $SE = 0.21$, $\beta = 0.15$, $t = 2.84$, $p = .005$) and AOT scores ($B = 0.26$, $SE = 0.10$, $\beta = 0.14$, $t = 2.62$, $p = .009$) significantly predicted overall BBS scores. As in Experiment 1, we correlated ICAR and AOT with the average assessments of self- and other-bias. Contrary to Experiment 1, ICAR was correlated with other-bias ratings ($r = .20$, $p < .001$), but not with self-bias ratings ($r = .01$, $p = .87$). Likewise, AOT significantly correlated with other-bias ratings ($r = .16$, $p = .002$) but not with self-bias ratings ($r = –.02$, $p = .69$).

As Table 5 shows, all six biases were observed in Experiment 2. Responses to the open-response anchoring item were Winsorized prior to analysis. Following West et al. (2012), to test the cognitive sophistication hypothesis, we conducted regression analyses where responses to bias items were predicted by the corresponding bias item set (A or B) and cognitive ability (ICAR). Critically, we included the interaction term, which provides a test of whether the effect of form varied as a function of cognitive ability. The critical set × ICAR interaction was not significant for any of the social/motivational items, nor was it significant for the conjunction bias item, all $p > .30$. However, this interaction was significant for the anchoring bias item ($B = −367.24$, $SE = 168.89$, $t = 2.17$, $p = .030$), the outcome bias item ($B = 1.65$, $SE = 0.49$, $t = 3.37$, $p = .001$), and the belief bias item ($B = 0.60$, $SE = 0.21$, $t = 2.83$, $p = .005$). The form of the interaction in all cases was such that,

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

as ICAR increased, the difference between the sets decreased (i.e., the magnitude of the bias decreased). With respect to the cognitive bias items, the two that Stanovich and West (2008) had shown were related to cognitive ability (i.e., outcome bias and belief bias) were also related to cognitive ability in Experiment 2. Of the two cognitive biases that Stanovich and West (2008) had shown were not related to cognitive ability, one was shown to not be related to cognitive ability in Experiment 2 (i.e., conjunction bias) but the other bias (i.e., anchoring) was related to cognitive ability. Therefore, we partially replicated the earlier results of Stanovich and West (2008).

Table 5: Responses to bias items in Experiment 2.

	Set A		Set B					
	M	*SD*	*M*	*SD*	*t*	*p*	*d*	95% CI
Self-Serving Bias	0.58	1.560	−0.72	1.723	−7.82	<.001	−0.79	−0.99, −0. 59
Halo Effect	4.05	0.830	3.75	1.238	−2.70	.007	−0.28	−0.48, −0.09
Outcome Bias	5.84	1.161	4.90	1.619	−6.73	<.001	−0.67	−0.87, −0.47
Belief Bias	1.71	0.498	0.43	0.667	−21.86	<.001	−2.17	−2.42, −1.92
Conj. Fallacy	43.75	27.27	29.29	23.86	−5.68	<.001	−0.56	−0.76, −0.36
Anchoring Bias	189.54	147.56	844.74	643.29	14.11	<.001	1.40	1.18, 1.62

Note. Experiment 2 ($N = 404$). All inferential statistics represent the results of independent-samples *t*-tests comparing the responses of participants presented with Set A to those presented with Set B.

Using the same method, we examined whether the effect of bias set varied as a function of AOT. The only bias for which the set × AOT interaction effect was significant was the outcome bias ($B = 0.54$, $SE = 0.24$, $\beta = 0.19$, $t = 2.23$, $p = .026$). The form of the interaction was similar to that observed with ICAR — as AOT increased, the effect of form was reduced.

Lastly, we applied a similar approach to assess whether the bias form interacted with the specific corresponding BBS (e.g., responses to the outcome bias item predicted by outcome bias blind spot) to predict responses to bias items. Only the interaction term for the halo effect was significant ($B = -0.14$, $SE = 0.06$, $t = 2.32$, $p = .021$). Here, the larger the halo bias blind spot, the smaller the halo bias. Contrary to the predictions of the cognitive sophistication hypothesis, all other interaction terms were non-significant, all $ps > .14$.

Using an alternative approach to test the cognitive sophistication hypothesis, we examined whether the relation between ICAR and individual BBS item scores was mediated by participants' degree of bias. We used the PROCESS module in SPSS (Hayes, 2018), and examined only those items for which a significant interaction effect between form and ICAR was identified in the previous section: outcome bias, belief bias, and anchoring bias. To construct a bias measure, we first created difference scores for each Bias item by subtracting individual scores in one bias form from the mean Bias score for the other form; that is, for

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

each individual who received Set A of the anchoring bias item, we subtracted the mean response of Set B. We then reversed the sign of one form for each item such that higher values consistently indicated greater bias. Finally, we standardized these values by calculating Z scores. We then ran a mediation analysis for each of the three BBS items, entering ICAR as a predictor and the relevant Z transformed biases as the mediating variable. Contrary to the predictions of the cognitive sophistication hypothesis, no significant mediation effect was identified for any of the individual items.

3.2.3 Bias blind spot strategies

To test the introspection bias hypothesis, we conducted a 2 (Bias Type: social/motivational, cognitive) × 2 (Target: self, other) × 2 (Strategy: introspection, behavior) repeated-measures ANOVA on strategy ratings (see Table 6 for a comprehensive set of pairwise test results). Recall that according to the introspection bias hypothesis, we would expect a significant target × strategy interaction effect such that introspection is the dominant strategy for self-bias ratings and behavior analysis is the dominant strategy for other-bias ratings. We observed a main effect of strategy ($F(1, 403) = 91.32$, $p < .001$, $\eta^2_p = .185$). Overall, participants reported using the behavior strategy ($M = 4.74$, $SD = 1.25$) more than the introspection strategy ($M = 4.36$, $SD = 1.35$).[7] Finally, there was a significant target × strategy interaction effect, $F(1, 403) = 8.53$, $p = .004$, $\eta^2_p = .021$. Specifically, while participants used the behavior strategy to a greater extent than the introspection strategy when assessing the degree of bias in both themselves ($t[403] = -5.82$, $p < .001$) and others ($t[403] = -10.46$, $p < .001$), this difference was larger when assessing the degree of bias in others ($d = -0.33$, 95% CI [−0.40, −0.27]) compared to when the target was themselves ($d = -0.21$, 95% CI [−0.28, −0.14]), $t(403) = 2.96$, $p = .003$, $d = 0.17$, 95% CI (0.06, 0.28). Other effects in the model were not statistically significant, all ps > .165.

Lastly, we computed a measure of Pronin and Kugler's (2007) introspection bias (i.e., a measure of the differential reliance on getting in the head (i.e., introspection) vs. using general behavior tendencies when judging oneself vs. others) as the difference between the use of the head and behavior strategy for the self and the other, respectively. As predicted by Pronin and Kugler (2007), this introspection bias score was correlated positively with overall BBS ($r = .11$, $p = .026$), although the correlation is small. Moreover, when ICAR, AOT and the introspection bias score were used to predict overall BBS, ICAR ($B = 0.06$, $SE = 0.02$, $\beta = 0.14$, $t = 2.72$, $p = .007$) and AOT ($B = 0.03$, $SE = 0.01$, $\beta = 0.14$, $t = 2.67$, $p = .008$) were significant predictors, but the introspection bias score was not significant ($B = 0.05$, $SE = 0.05$, $\beta = 0.05$, $t = 0.98$, $p = .329$).

[7]We observed a main effect of bias type ($F[1, 403] = 16.71$, $p < .001$, $\eta^2_p = .040$), but this was not of theoretical interest and already examined in an ANOVA reported earlier.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

TABLE 6: Bias blind spot introspection and behavior strategy ratings in Experiment 2.

		Other		Self					
Bias Type	Strategy	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*	95% CI
All biases	Intro.	4.35	1.40	4.38	1.46	−0.61	.543	−0.02	−0.09, 0.04
	Beh.	4.80	1.29	4.67	1.35	3.07	.002	0.10	0.04, 0.16
Soc./Mot.	Intro.	4.41	1.56	4.45	1.56	−0.74	.457	−0.03	−0.11, 0.05
	Beh.	4.89	1.40	4.78	1.49	2.10	.037	0.08	0.00, 0.15
Halo	Intro.	4.36	1.70	4.43	1.69	−0.94	.346	−0.04	−0.13, 0.04
	Beh.	4.82	1.56	4.74	1.66	1.16	.249	0.05	−0.03, 0.13
Self-Serving	Intro.	4.45	1.68	4.47	1.68	−0.30	.765	−0.01	−0.10, 0.07
	Beh.	4.96	1.53	4.81	1.62	2.25	.025	0.10	0.01, 0.17
All Cognitive	Intro.	4.32	1.42	4.34	1.48	−0.39	.693	−0.01	−0.08, 0.05
	Beh.	4.75	1.34	4.62	1.38	2.86	.004	0.10	0.03, 0.17
Cog.-Related	Intro.	4.33	1.51	4.33	1.59	−0.39	.693	0.00	−0.08, 0.08
	Beh.	4.89	1.43	4.70	1.53	2.86	.004	0.13	0.05, 0.21
Outcome	Intro.	4.40	1.63	4.42	1.70	−0.38	.705	−0.02	−0.10, 0.07
	Beh.	5.07	1.53	4.82	1.64	3.46	.001	0.15	0.07, 0.24
Belief	Intro.	4.27	1.70	4.24	1.75	0.31	.756	0.01	−0.08, 0.10
	Beh.	4.71	1.66	4.57	1.70	1.76	.080	0.08	−0.01, 0.17
Cog.-Unrelated	Intro.	4.31	1.51	4.35	1.56	−0.62	.534	−0.03	−0.11, 0.05
	Beh.	4.62	1.45	4.53	1.45	1.45	.149	0.06	−0.02, 0.14
Conjunction	Intro.	4.34	1.59	4.45	1.66	−1.39	.164	−0.06	−0.15, 0.03
	Beh.	4.74	1.54	4.61	1.60	1.89	.060	0.09	−0.00, 0.17
Anchoring	Intro.	4.27	1.66	4.25	1.73	0.32	.752	0.01	−0.08, 0.11
	Beh.	4.50	1.65	4.46	1.66	0.50	.617	0.02	−0.07, 0.12

Note. *df* = 403. All inferential statistics are based on paired-samples *t*-tests comparing ratings provided across self and other items. Soc./Mot. = Social/Motivational; Cog. = Cognitive; Intro. = introspection; Beh. = behavior.

4 General discussion

Across a wide range of social/motivational and cognitive biases, individuals more readily judge that these biases afflict others more strongly than they afflict themselves. This BBS was evident across several distinct biases in our research, and in both experiments, the various bias blind spots were intercorrelated. Moreover, in Experiment 2 where we tested single and multi-component models of BBS using principal components analysis, the evidence did not justify the use of more than a single component, even though we used a mix of

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

social/motivational and cognitive bias items. Finally, support for the BBS in Experiment 2 was found using a novel randomized-block design that included a much greater degree of self/other rating separability while retaining the power of a repeated-measures design.

4.1 Testing explanatory hypotheses

Consistent with West et al. (2012), we found that the BBS positively correlated with measures of cognitive sophistication in both experiments. Individuals who scored higher on a measure of cognitive ability and a measure of actively open-minded thinking also registered larger bias blind spots. Going beyond West et al. (2012), we demonstrated this relation across several biases that were not previously examined, including social/motivational biases, and also found that both cognitive ability and thinking dispositions made unique contributions in predicting the BBS. Moreover, we established these relations using an objective measure of cognitive ability, whereas earlier studies had typically used self-reports of SAT scores, which may be prone to errors or systematic distortions. The fact that cognitive ability and thinking dispositions uniquely predict the BBS is consistent with previous research showing their independent contributions to reasoning performance and that thinking dispositions often influence reasoning performance over and above cognitive ability (e.g., Stanovich & West, 1997, 1998). These findings, therefore, contribute to the literature on the joint contributions of cognitive ability and thinking style to explain reasoning and decision-making performance (Evans & Stanovich, 2013).

However, our findings also revealed a surprising result regarding relations between the BBS and measures of cognitive sophistication. In Experiment 1, these relations were driven almost entirely by negative correlations between the self-bias ratings and the cognitive sophistication measures (i.e., ICAR and AOT). In contrast, in Experiment 2, these relations were driven almost entirely by positive correlations between other-bias ratings and the cognitive sophistication measures. In different terms, cognitive sophistication was associated with judging oneself to be less biased in Experiment 1, whereas it was associated with judging others to be more biased in Experiment 2. These differences are masked by the fact that the BBS measures represent a difference score, namely, a linear combination of two primary dependent measures. Given these results, we find it unsurprising that the evidence for the relation between the BBS and cognitive sophistication measures has been mixed in the literature. We recommend that future studies exploring such relations examine the primary dependent measures separately in addition to the composite BBS measure.

In line with West et al. (2012), we found limited evidence in Experiment 2 that individual differences in the BBS and its relations with cognitive ability and AOT are due to lower susceptibility to the actual biases. Based on the present findings, then, we must reject the cognitive sophistication hypothesis. Although the BBS appears to be positively related to cognitive sophistication, this relation was not mediated by actual susceptibility to biases. Moreover, we emphasize that the observed relations between the BBS and measures of cognitive sophistication, even when significant, have been invariably small in magnitude.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

In Experiment 2, we also set out to replicate and extend Pronin and Kugler's (2007) demonstration that participants rely on different strategies when making self versus other judgments of bias susceptibility. The results here were mixed. Pronin and Kugler (2007) found a crossover pattern wherein participants reported relying more on "getting inside one's head" (i.e., introspection) than "general behavioral tendencies" when judging themselves and the opposite tendency when judging others. We did not find this crossover pattern. Participants in Experiment 2 tended to report relying more on general behavioral tendencies than on getting inside one's head. However, the magnitude of this difference was smaller when individuals were assessing themselves compared to assessing others. This relative difference is consistent with the introspection bias hypothesis, although as noted, the characteristics of the interaction fall short of the specific prediction of the hypothesis. Moreover, the interaction effect size was small and may represent no more than an inherent constraint on introspection, in general. It is debatable whether one could literally introspect for another individual. Indeed, Ehrlinger et al. (2005) make the point that "people cannot, by definition, introspect into the minds of others" (p. 681). As well, we found that participants' introspection bias scores did not predict the BBS when controlling for cognitive sophistication measures. In fact, even when cognitive sophistication was not controlled, the correlation between introspection bias and the BBS explained approximately 1% of the variance. Therefore, while we found some support for the introspection bias hypothesis, we believe our results call into question whether introspection bias constitutes a principal mechanism for explaining variation in people's bias blind spots. We believe further research using a variety of methodological approaches will be required to settle this question.

A third and novel explanation that we tested in our research is that the BBS is largely an artifact of the conversational logic of the task. In the typical task, participants are told before making their assessments that the biases they are evaluating are ones that are exhibited by *people in general*, and participants have no reason to doubt that these findings also apply to the *average* survey respondent (Hilton, 1995; Sperber & Wilson, 1986). Accordingly, one might expect to find a greater proportion of participants who provided ratings for others that fell above the midpoint of the rating scale, as that pattern of response would be tantamount to repeating what one had just been told was true, in general. In fact, this is what we observed in both experiments. The proportion of participants giving other-bias assessments above the scale midpoint was significantly greater than the proportion of participants giving self-bias ratings above the midpoint. Moreover, the proportion of participants giving such ratings for others, in general, represented significant majorities in both experiments. Finally, consistent with the conversational processes hypothesis, we found in both experiments that the self-bias ratings were invariably more variable than the other-bias ratings, which we would expect if the stated scientific facts have more bearing on "respondents, in general" than "me (the participant), in particular."

The conversational hypothesis is supported by at least one other study of the BBS. In their first study, Pronin et al. (2002) asked participants to rate their own susceptibility to

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

biases as well as the susceptibilities of the average American and one of their parents. They found that although the self and parent ratings indicated lower bias susceptibility than the average-American ratings, there was no significant difference between the self and parent ratings, except on one item (positive halo effect) in which case the parent-rating was significantly *lower* than the self-rating. In other words, there was no tendency to exhibit bias blind spots when comparing oneself to a specific other about which one has ample individuating information and is less likely to be constrained by information about people, in general.

The conversational hypothesis could be profitably examined in future research. For example, aside from strictly conversational processes, we might expect differences in self and other ratings because self-ratings will mainly engage singular or "inside view" judgment processes, whereas asking about an average other will mainly engage distributional or "outside view" judgment processes (Kahneman, 2011; Reeves & Lockhart, 1993). Instead of asking solely about the average survey respondent (or the average other in any other cohort), researchers could ask about specific others that participants know well (e.g., a close friend, romantic partner, or family member) or do not know well (e.g., an acquaintance). Future studies might also draw on findings from the literature on unrealistic optimism. For instance, Klar et al. (1996) found no significant difference in unrealistic optimism for *controllable* events between assessments of self and a specific close other. However, both self and specific close others were assessed more optimistically than the *average* peer. For *uncontrollable* events, none of the targets significantly differed. Future studies of the BBS could likewise attempt to manipulate how controllable the biases appear to be given that controllability influences self-other comparisons underlying unrealistic optimism. This could potentially be accomplished, for instance, by describing to participants either the successes or the failures of interventions aimed at debiasing the same types of biases they are asked to assess.

Investigating the cause of the BBS is important. The literature currently tends to assume that the BBS is a metacognitive error in which people underweight their relative contribution to psychological biases (e.g., Jones et al., 2018). Unsurprisingly, then, recent research has focused on developing training interventions to mitigate the bias blind spot (Bessarabova et al., 2016). However, if the cause or causes of this differential response pattern to self and other assessments is uncertain, it also casts doubt on the requirement for prescriptive interventions. If that pattern is attributable mainly to reasonable conversational inferences that participants draw in BBS experiments, then the BBS will have little implication for pernicious behavioral and social effects in the real world and such mitigation efforts would be well directed elsewhere.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

References

Arkes, H. R. (2013). The consequences of the hindsight bias in medical decision making. *Current Directions in Psychological Science*, *22*(5), 356–360.

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, *4*(3), 265–284.

Bessarabova, E., Piercy, C. W., King, S., Vincent, C., Dunbar, N. E., et al. (2016). Mitigating bias blind spot via a serious video game. *Computers in Human Behavior, 62,* 452–66.

Bornstein, B. H., & Emler, A. C. (2001). Rationality in medical decision making: A review of the literature on doctors' decision-making biases. *Journal of Evaluation in Clinical Practice*, *7*(2), 97–107.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*(2), 197–253.

Chandrashekar, S. P., Yeung, S. K., Yau, K. C., Cheung, C. Y., Agarwal, T. K., Wong, C. Y. J., Pillai, T., Thirlwell, T. N., Leung, W. N., Tse, C., Li, Y. T., Cheng, B. L., Chan, H. Y. C., & Feldman, G. (2021). Agency and self-other asymmetries in perceived bias and shortcomings: Replications of the bias blind spot and link to free will beliefs. *Judgment and Decision Making*, *16*, 1392–1413.

Condon, D.M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence, 43*, 52–64.

Dhami, M. K., Belton, I. K., & Mandel, D. R. (2019). The "analysis of competing hypotheses" in intelligence analysis. *Applied Cognitive Psychology*, *33*(6), 1080–1090.

Dworak, E. M., Revelle, W., Doebler, P., Condon, D. M. (2021). Using the international cognitive ability resource as an open tool to explore individual differences in cognitive ability. *Personality and Individual Differences, 169*, 1–9.

Ehrlinger, J, Gilovich, T., & Ross, L. (2005). Peering into the bias blind spot: people's assessments of bias in themselves and others. *Personality and Social Psychology Bulletin, 31*(5), 680–692.

Elashi, F. B., & Mills, C. M. (2015). Developing the bias blind spot: Increasing skepticism towards others. *PLoS ONE, 10*(11): e0141809. https://doi.org/10.1371/journal.pone.0141809

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.) (2002). *Heuristics and biases: The psychology of intuitive judgement*. Cambridge University Press.

Hagá, S., Olson, K. R., & Garcia-Marques, L. (2018). The bias blind spot across childhood. *Social Cognition, 36*(6), 671–708.

Haran, U., Ritov, I., & Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Judgment and Decision Making, 8*(3), 188–201.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd edition). The Guilford Press.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin, 118*, 248–271.

Hooper, D., Couglan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60.

Jones, K. A., Crozier, W. E., & Strange, D. (2018). Objectivity is a myth for you but not for me or police: A bias blind spot for viewing and remembering criminal events. *Psychology, Public Policy, and Law*, *24*(2), 259–270.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus & Giroux.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, *2*(1), 42–52.

Klar, Y., Medding, A., & Sarel, D. (1996). Nonunique invulnerability: Singular versus distributional probabilities and unrealistic optimism in comparative risk judgments. *Organizational Behavior and Human Decision Processes, 67,* 229–245.

Kline, R. B. (2005). *Methodology in the social sciences. Principles and practice of structural equation modeling (2nd edition).* Guilford Press.

Kukucka, J., Kassin, S. M., Zapf, P. A., & Dror, I. E. (2017).Cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory and Cognition*, *6*(4), 452–459.

Littrell, S., & Fugelsang, J. A. (2021). *The 'bullshit blind spot': The roles of overconfidence and perceived information processing in bullshit detection*. Manuscript submitted for publication. *PsyArxiv,* https://psyarxiv.com/kbfrz/

Mandel, D. R., Collins, R. N., Risko, E. F., & Fugelsang, J. A. (2020). Effect of confidence interval construction on judgment accuracy. *Judgment and Decision Making, 15*(5), 783–797.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Erlbaum.

Pronin, E., Berger, J. & Molouki, S. (2007). Alone in a crowd of sheep: Asymmetric perceptions of conformity and their roots in an introspection illusion. *Journal of Personality and Social Psychology, 92*(4), 585-595.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*(3), 369–381.

Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology*, *43*(4), 565-578.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, *111*(3), 781–799.

Raoelison, M., Boissin, E., Borst, G., & De Neys, W. (2021). From slow to fast logic: The development of logical intuitions. *Thinking & Reasoning*, *27*(4), 599–622.

Reeves, T., & Lockhart, R. (1993). Distributional versus singular approaches to probability and errors in probabilistic reasoning. *Journal of Experimental Psychology: General, 122*, 207 –226.

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment*, *31*(12), 1395–1411.

Scopelliti, I., Morewedge, C. K., McCormick, E., Min, H. L., Lebrecht, S., & Kassam, K. S. (2015). Bias blind spot: Structure, measurement, and consequences. *Management Science*, *61*(10), 2468–2486.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Blackwell.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*(2), 342–357.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127(2),* 161–188.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*(4), 672–695.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*, 1275–1289.

Wen, W., Li, J., Georgiou, G. K., Huang, C., & Wang, L. (2020). Reducing the halo effect by stimulating analytic thinking. *Social Psychology, 51*, 334–340.

West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of Personality and Social Psychology*, *103*(3), 506–519.

Zappala, M., Reed, A. L., Beltrani, A., Zapf, P. A., & Otto, R. K. (2018). Anything you can do, I can do better: Bias awareness in forensic evaluators. *Journal of Forensic Psychology Research and Practice*, *18*(1), 45–56.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Appendix A: BBS Items in Experiment 1

Items (All Biases): To what extent do you believe that you show this effect or tendency?

To what extent do you believe the average survey respondent shows this effect or tendency?

Response Options: 1 (Not At All); 2; 3; 4; 5; 6; 7 (Very Much)

Action-Inaction Bias

Some people show a tendency to judge a harmful action as worse than an equally harmful inaction. For example, this tendency leads to thinking it is worse to falsely testify in court that someone is guilty, than not to testify that someone is innocent.

Bandwagon Effect

Psychologists have claimed that some people show a tendency to do or believe a thing only because many other people believe or do that thing, to feel safer or to avoid conflict.

Confirmation Bias

Many psychological studies have shown that people react to counterevidence by actually strengthening their beliefs. For example, when exposed to negative evidence about their favorite political candidate, people tend to implicitly counterargue against that evidence, therefore strengthening their favorable feelings toward the candidate.

Disconfirmation Bias

Psychologists have claimed that some people show a "disconfirmation" tendency in the way they evaluate research about potentially dangerous habits. That is, they are more critical and skeptical in evaluating evidence that an activity is dangerous when they engage in that activity than when they do not.

Diffusion of Responsibility

Psychologists have identified an effect called "diffusion of responsibility," where people tend not to help in an emergency situation when other people are present. This happens because as the number of bystanders increases, a bystander who sees other people standing around is less likely to interpret the incident as a problem, and also is less likely to feel individually responsible for taking action.

Escalation of Commitment

Research has found that people will make irrational decisions to justify actions they have already taken. For example, when two people engage in a bidding war for an object, they can

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

end up paying much more than the object is worth to justify the initial expenses associated with bidding.

Fundamental Attribution Error

Psychologists have claimed that some people show a tendency to make "overly dispositional inferences" in the way they view victims of assault crimes. That is, they are overly inclined to view the victim's plight as one he or she brought on by carelessness, foolishness, misbehavior, or naivetë.

Halo Effect

Psychologists have claimed that some people show a "halo" effect in the way they form impressions of attractive people. For instance, when it comes to assessing how nice, interesting, or able someone is, people tend to judge an attractive person more positively than he or she deserves.

Ingroup Favoritism

Extensive psychological research has shown that people possess an unconscious, automatic tendency to be less generous to people of a different race than to people of their race. This tendency has been shown to affect the behavior of everyone from doctors to taxi drivers.

Ostrich Effect

Psychologists have identified a tendency called the "ostrich effect," an aversion to learning about potential losses. For example, people may try to avoid bad news by ignoring it. The name comes from the common (but false) legend that ostriches bury their heads in the sand to avoid danger.

Projection Bias

Many psychological studies have found that people have the tendency to underestimate the impact or the strength of another person's feelings. For example, people who have not been victims of discrimination do not really understand a victim's social suffering and the emotional effects of discrimination.

Self-Interest Bias

Psychologists have claimed that some people show a "self-interest" effect in the way they view political candidates. That is, people's assessments of qualifications, and their judgments about the extent to which particular candidates would pursue policies good for the American people as a whole, are influenced by their feelings about whether the candidates' policies would serve their own particular interests.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Self-Serving Bias

Psychologists have claimed that some people show a "self-serving" tendency in the way they view their academic or job performance. That is, they tend to take credit for success but deny responsibility for failure. They see their successes as the result of personal qualities, like drive or ability, but their failures as the result of external factors, like unreasonable work requirements or inadequate instructions.

Stereotyping

Psychologists have argued that gender biases lead people to associate men with technology and women with housework.

Appendix B: BBS Items in Experiment 2

Social Biases (adapted from Scopelliti et al., 2015)

Halo Effect

Psychologists have claimed that some people show a "halo" effect in the way they form impressions. That is, they allow their knowledge of one characteristic of an individual to influence their judgment of other characteristics (even if those characteristics are unrelated). For example, when it comes to assessing how nice, interesting, or able someone is, people tend to judge an attractive person more positively than he or she deserves.

Self-Serving Bias

Psychologists have claimed that people show a "self-serving" tendency in the way they view their academic or job performance. That is, they tend to take credit for success but deny responsibility for failure; they see their successes as the result of personal qualities, like drive or ability, but their failures as the result of external factors, like unreasonable work requirements or inadequate instruction. For example, if a car dealer met her sales goals for the year, she may attribute it to her being a great salesperson. However, if she did not meet her sales goals, she may blame the economy.

Cognitive Biases Related to Cognitive Ability (Outcome Bias adapted from West et al., 2012)

Outcome Bias

Psychologists have found that people tend to judge the quality of a decision based on how the decision worked out. That is, people sometimes forget that the quality of the decision must be judged on what was known at the time the decision was made, not how

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

it worked out, because the outcome is not known at the time of the decision. For example, if a weatherperson reports a 75% chance of rain and as a result tells viewers to bring an umbrella if they go outside, people will often praise the weatherperson's advice if it rains, but criticize the weatherperson's advice if it does not even though it was arguably the best advice to give at the time.

Belief Bias

Psychologists have found that people's beliefs can influence the evaluation of logical arguments. This bias is often assessed using problems where the believability of the conclusion conflicts with the logical validity. Here, people tend to accept as valid conclusions that are consistent with their personal beliefs even if they do not logically follow from the premises. For example, when given the syllogism "all vehicles have wheels, a boat is a vehicle, therefore, a boat has wheels", many people will judge this conclusion as invalid, by relying on their prior beliefs about boats, despite it actually being a logically valid conclusion given the information in the premises.

Cognitive Biases Unrelated to Cognitive Ability (Outcome Bias adapted from West et al., 2012)

Conjunction Fallacy

Psychologists have found that people tend to rate conjunctions of events (situations where two or more truth conditions must be met) as too likely. Conjunctions of events become less likely as the number of truth conditions grow. For example, suppose (A) 90% of the objects in a jar are marbles and (B) 90% of objects in the jar are red, people may be prone to assign a higher probability to the probability of drawing a red marble at random from the jar than simply drawing a marble at random, even though "red marbles" are a subset of "marbles" and therefore cannot be more likely.

Anchoring

Psychologists have found that people making numerical estimations tend to focus on any number that is available to help them. This is a good strategy, except in situations where the available numbers are unrelated to the quantity we are trying to estimate. For example, people report fewer headaches when they are asked: "How many headaches do you have a month — 0, 1, 2 — how many?" than when they are asked: "How many headaches do you have a month — 5, 10, 15 — how many?"

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Appendix C: Cognitive Bias Tasks in Experiment 2

Self-serving bias

Participants were provided with one of the following questions, for which experience type (positive or negative) was manipulated across sets.

Set A (Positive Experience): Think of a time when you experienced something really positive. To what extent was this due to things you did or to factors beyond your control?

Set B (Negative Experience): Think of a time when you experienced something really negative. To what extent was this due to things you did or to factors beyond your control?

Participants responded to this item on a 7-point scale that ranged from −3 (*totally beyond my control*) to 3 (*totally within my control*). Self-serving bias is observed if higher ratings are provided by participants asked to think about a positive (Set A) as opposed tonegative (Set B) experience.

Halo effect

Participants were given a description of a hypothetical agent (i.e., Professor T) and asking them to rate their agreement with the statement, "Professor T is good-looking." This item was based on an item used in Wen et al. (2020). Ratings were provided on a 7-point scale that ranged from *strongly disagree* to *strongly agree*. Whether Professor T was described as treating students coldly or warmly was manipulated across forms (see below).

Set A (Treats Students Coldly): Professor T is a college teacher who works hard, solves problems quickly, and treats students coldly.

Set B (Treats Students Warmly): Professor T is a college teacher who works hard, solves problems quickly, and treats students warmly.

A halo effect is observed if participants agree more with the statement that Professor T is good-looking when Professor T is described as treating students warmly (Set B) as opposed to coldly (Set A).

Anchoring bias

The anchoring bias item was similar to that used in Experiment 2, with the exception that participants now answered a single problem in which they were asked to estimate "how tall do you think the tallest redwood tree in the word is (in feet)?" after first being asked to indicate whether this tree was taller or shorter than either a small (85 feet) or large (1,000 feet) anchor. The size of the anchor (i.e., 85 feet or 1,000 feet) was manipulated across forms. Anchoring bias is observed if participants provide larger estimates when provided with the large as opposed to small anchor.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press

Belief bias

Participants were presented with two syllogisms (one valid and one invalid), taken from Raoelison et al., (2021). The believability of conclusions was manipulated across sets.

Set A (Believable/Invalid): Everything with a motor needs oil. Cars need oil. Therefore, cars have a motor.

Set A (Believable/Valid): All birds have wings. Crows are birds. Therefore, crows have wings.

Set B (Unbelievable/Invalid): All African countries are warm countries. Spain is a warm country. Therefore, Spain is an African country.

Set B (Unbelievable/Valid): All mammals can walk. Whales are mammals. Therefore, whales can walk.

Each syllogism was presented such that the first premise, second premise, 0and conclusion were clearly indicated. For each syllogism, participants were asked, "Does the conclusion follow logically?" and they responded with either a "Yes" or "No" response. Belief bias is observed if participants evaluating syllogisms with believable conclusions (Set A) respond "Yes" more often than those evaluating syllogisms with unbelievable conclusions (Set B). As such, "Yes" responses were summed to create a belief bias score for each participant that ranged from 0 to 2.

https://doi.org/10.1017/S1930297500009475 Published online by Cambridge University Press