

The Privileged Information Effect

Stephen M. Baum¹ and Clayton R. Critcher¹

¹ Haas School of Business, University of California, Berkeley

Author Note

Stephen M. Baum and Clayton R. Critcher, Haas School of Business, University of California, Berkeley.

This work was supported in part by U.S. National Science Foundation award 1749608 awarded to Clayton R. Critcher.

Correspondence concerning this article should be addressed to Stephen M. Baum, Haas School of Business, University of California, Berkeley, 545 Student Services Building #1900, Berkeley, CA 94720. E-mail: stephen_baum@berkeley.edu

Abstract

Different people learn the same information at different times. As a result, information asymmetries abound in marketplaces, workplaces, and close relationships. Such asymmetries are often temporary: What is privileged information today may be common knowledge tomorrow. For example, one hiring committee member may receive a newly submitted recommendation letter a day before other committee members. Previous research has documented challenges with reasoning across such open information asymmetries. This research instead examines the anticipated consequences of closing information asymmetries. Seven experiments document a *privileged information effect* (PIE): Privileged information holders think information about a target—positive or negative—will more strongly influence others' evaluations upon learning it as compared to if others had known it all along. This PIE emerged across diverse contexts (e.g., product reviews, hiring recommendations, romantic matchmaking) and reflected a bias. Three possible mechanistic accounts were identified; only a *segregated processing account* was supported. Privileged information holders naturally segregate what another knows from what they *could* know (but is currently privileged). Thus, when privileged information holders add this (currently privileged) information into their understanding of another's evaluation, it is not incorporated on equal footing with what is common knowledge. In the absence of an intervention that discourages this segregated processing, the privileged information is heavily weighted in estimates of others' evaluations. Alternative accounts rooted in focalism and the perception that secrets have more evaluatively extreme meaning were empirically considered but not supported.

KEYWORDS: information asymmetry, privileged information, social judgment, secrets

The Privileged Information Effect

Consider Josh and Colleen, members of a three-person hiring committee that is in the process of evaluating Sally, a promising job candidate. Certain positive aspects of Sally's competence and character—such as her educational background, employment history, and qualities expanded upon in her letters of recommendation—would be known by the entire committee. Yet Josh may have additional information about Sally that is not yet common knowledge—for example, a negative detail shared in a reference letter sent only to him and Colleen. This produces an *information asymmetry*, a circumstance in which one person is aware of information that another is not.

Asymmetries in experience give rise to asymmetries in knowledge. People's experiences vary, so information asymmetries abound in markets (Akerlof, 1970; Mishra, Heide, & Gort, 1998), negotiations (Thaler, 1988), close relationships (Friedberg & Stern, 2014) and, indeed, hiring decisions (Desai, Lockett, & Paton, 2017). From the letter that not everyone has seen, Josh might know that Sally can be quite arrogant, often promoting herself as more talented than her coworkers. Certainly, such *privileged information* would color Josh's impressions of Sally: He will evaluate her candidacy less positively having learned this unflattering detail. But how would it—and its *privileged* status in particular—influence Josh's estimates about what the other members of the committee think of her?

Estimating others' evaluations often entails considering not just what others *currently* know, but what others will *come to know*. While Josh and Colleen have the same information about Sally's candidacy, Josh and the third committee member—who was not privy to the additional reference letter—do not. Thus, Josh might consider how this third committee member would evaluate Sally if they actually did learn about her arrogance. After all, privileged

information—whether in the workplace, the classroom, or the community—often becomes common knowledge eventually (Slepian, Chen, & Mason, 2017; Slepian & Greenaway, 2018).

In this manuscript, we test whether information that is not yet (vs. already) known by another is seen as particularly likely to influence their overall evaluations of a target. In the context of our opening example, Josh may believe that the third committee member—if they *were* to receive the privileged information—would evaluate Sally’s candidacy more negatively than does Colleen (who learned of Sally’s arrogance when Josh did). We call this proposed phenomenon the *privileged information effect* (PIE). Ultimately, we will demonstrate the PIE in a variety of information asymmetry contexts (e.g., in considering how another does (or will) evaluate a new technology product, a job candidate, and even love interests). These studies will establish the effect’s robustness (e.g., that the PIE emerges both when privileged information is positive or negative), consider the extent to which the PIE is a social bias (i.e., whether initially privileged information, once learned, actually does exert as much of an influence on evaluations as expected), and distinguish among multiple mechanisms to pinpoint what gives rise to the PIE.

Information Asymmetries and Social Judgments

To return to our opening example, several literatures in psychology and economics suggest that Josh would have difficulty ignoring privileged information when estimating another’s perspective. Most notably, research on the *curse of knowledge* illustrates how better-informed parties have difficulty setting aside their privileged information to appreciate the perspective of lesser-informed others (Camerer, Loewenstein, & Weber, 1989; Chambers, Epley, Savitsky, & Windschitl, 2004; Keysar, Ginzler, & Bazerman, 1995). A former employee who suffered years of verbal abuse from an overbearing boss may be surprised to learn just how likeable the boss’s outside-of-work acquaintances find him to be. This curse of knowledge is one

of many phenomena that reflect the egocentrism that characterizes much of social thinking. Furthermore, the self's difficulty with dissociating its own knowledge from what it believes others know also extends to the self's knowledge of its own internal states. People show an *illusion of transparency*, overestimating the extent to which their own privileged internal states are apparent to others (Gilovich, Savitsky, & Medvec, 1998). People even have trouble setting aside their present perspective in remembering that of their previous self: Because it is difficult to “unknow” information once acquired, people display a hindsight bias in remembering just how ignorant their past self was (Fischhoff, 1975; cf. Dietvorst & Simonsohn, 2019).

Such egocentrism can plague the lesser informed as well. For example, Bazerman and Carrol (1987) put participants through a negotiation in which buyers—as they typically do—had less information about a company's value than did sellers. Buyers did not navigate their state of ignorance well. They ended up arriving at unprofitable deals, failing to appreciate that the better-informed sellers could easily exploit their own superior knowledge. Even when they *know* they operate with less knowledge than another, people still have trouble correcting for their own perspective in order to appreciate how another's is different (Samuelson & Bazerman, 1985). The potential for exploitation results. In short, the self's own perspective exerts inappropriate weight in reasoning about those with either less or more knowledge.

More broadly, asking how people understand the perspectives and attitudes of those who have more or less information than the self is a theory-of-mind question (Bretherton & Beeghly, 1982). In addressing such questions, social and personality psychologists have extensively examined *metaperceptions* - beliefs about how someone else views another (Kenny & DePaulo, 1993; Laing, Phillipson, & Lee, 1966). The most well-developed research in this tradition has examined people's beliefs about how others evaluate the self (Ambady & Rosenthal, 1992;

Kenny & Albright, 1987; Kenny, Bond, Mohr, & Horn, 1996). After all, beliefs about how one comes across to others shape how one navigates one’s personal and professional life (Donnelly, Moon, & Critcher, 2022). To move closer to our opening example, a smaller literature on *triadic perceptions*—also called *three-party metaperceptions* (Kenny & DePaulo, 1993; Kenny, 1994; Bond, Horn, & Kenny, 1997)—has examined one person’s estimate of how a second person feels about a third person. However, the present paper’s interests are a bit broader: We examine beliefs about how another person evaluates any target, not merely social ones (i.e., another person) but non-social targets as well (e.g., an investment opportunity, a product). For that reason, we introduce the phrase *social evaluative estimates* (SEEs) to describe beliefs about how another evaluates a target, either a person or an object.¹

In the context of information asymmetries, existing research has explored what influences or distorts SEEs made across *open* information asymmetries (i.e., situations in which one party is better-informed than another). Instead of examining how people reason about information asymmetries that are currently open, we will study how people reason about information asymmetries that *will* close (vs. do not exist). A job candidate may wonder how a hiring committee member would respond to learning from his background check that he recently got a DUI. A matchmaker may consider how a friend would evaluate a blind date upon learning that person has an amazing apartment. In both cases, one can consider others’ evaluations in prospect (“How would the hiring committee member evaluate me once they do receive the background check?”), but also when the information asymmetry is known to be closed (“How does the hiring committee member evaluate me, having now reviewed my entire application, which included the

¹ The *social* in *social evaluative estimates* denotes that the belief in question is about *another person’s* evaluation of a target. Thus, although the ultimate target of evaluation may be either a person or an object, SEEs refer to one person’s belief about another person’s evaluation of that target.

background check?”). When people are naturally curious about another’s current or potential private evaluations, these SEEs can easily take either form.

The pervasiveness of egocentrism might suggest it would not matter whether one considers someone who already has or someone who potentially will have the same information as the self does currently. After all, egocentrism misleads when the self’s own perspective is shaped by having more (or less) information than does another person. But with an information asymmetry that is (potentially) closing—or that does not exist—the self is considering another who possesses the same information. To return to our opening example, Josh could estimate Collen’s present, fully informed opinion of Sally or the third committee member’s potential, fully informed opinion of Sally by consulting his own fully informed perspective. In this way, egocentrism—the psychological mechanism most frequently implicated in the information asymmetry literature—would push these two SEEs to be identical and, beyond that, to be reasonably accurate (Dawes, 1989; Hoch, 1987). But as we will develop rhetorically and empirically below, we expect Josh’s estimate of what the third committee member will think of Sally to be more negative than his estimate of what Colleen already thinks of Sally.

Segregated Processing and the Privileged Information Effect

Our theoretical account begins with the fact that privileged information is not processed or incorporated into an SEE on equal footing with the broader set of facts understood to be common knowledge. Suppose that the information that all committee members already have about Sally can be consolidated into three positive characteristics: P1, P2, and P3. In contrast, the negative detail about Sally’s arrogance is known by Colleen (N1, common knowledge) but not by the third committee member (N1_{priv}, privileged information). In the absence of an information asymmetry, Josh aims to estimate another’s (i.e., Colleen’s) evaluations of a target defined by the

entire informational set (P1, P2, P3, N1). Here, N1—the negative detail—is merely one fact among (and incorporated into the SEE alongside) many attributes.

But when the information asymmetry is currently open, what is common knowledge (P1, P2, and P3) and what is privileged information ($N1_{priv}$) become naturally segregated. In this case, Josh is asking how someone (e.g., the third committee member) who has limited information (P1, P2, P3) will respond *upon* learning $N1_{priv}$. In turn, $N1_{priv}$ is not processed and incorporated on equal footing with the other attributes; rather, it is incorporated as a more isolated, standalone input. Of course, given information is typically encountered sequentially, *all* pieces of information are—in some sense—segregated from the last (and the rest). But we note that privileged information holders are in a unique position that reinforces the segregation in real time: They can simultaneously consider what someone else knows (e.g., P1, P2, P3) and what they could know (e.g., [P1, P2, P3] + N1). We argue that the privileged status afforded by this segregated processing leads privileged information to be particularly influential in forming such SEEs, thereby giving rise to the PIE.

Before proceeding, we acknowledge that this segregated processing account is related to—but distinct from—previously articulated *focalism* accounts. Focusing on a single piece of information can lead that information to exert a disproportionate influence on judgment (Higgins, 1996; Wilson et al., 2000). Information can become focal for a variety of reasons: when it is presented first or last in a sequence (Asch, 1946; O'Brien & Ellsworth, 2012), is easy to bring to mind (Gilbert, 1991; Kahneman & Miller, 1986; Roese & Olson, 1997; Shedler & Manis, 1986), is potentially threatening to the self (Critcher & Dunning, 2015; Moon, Gan, & Critcher, 2020), or is ruminated upon given the inexplicability of why it pops into mind (Morewedge, Giblin, & Norton, 2014). Of greater relevance to the present phenomenon, secrets

can become focal and dominate secret-holders' attention (Slepian, Chun, & Mason, 2017; Slepian, Kirby, & Kalokerinos, 2020).

But we are not arguing that privileged information—merely by dint of receiving special consideration among one's set of target-relevant information—will be top-of-mind and influence *all* of one's judgments. Instead, we note that when privileged information holders form SEEs, what *might* be known by another (privileged information) is added to an already formed SEE (that consists of what is *currently* known by another). This leads the privileged information to no longer be incorporated on equal, but instead on special (or privileged), footing, thereby encouraging such information to be weighted more heavily in SEEs. In what follows, we elaborate on this logic, identify four predictions that follow from this reasoning, and differentiate this segregated processing from alternative mechanisms (e.g., focalism).

First, the segregated processing account anticipates that the PIE should apply to SEEs in particular, not simply to any judgment that the privileged information holder makes (e.g., their *own* evaluations of a target). Privileged information holders naturally consider two SEEs simultaneously—one based on what the other *currently* knows (e.g., $P1 + P2 + P3$) and one based on what the other *might* come to know (i.e., the current SEE + $N1$). When estimating the third committee member's evaluation of Sally, Josh will naturally consider how this other will evaluate Sally both based on the set of features that are already common knowledge *and* how that SEE would change if the other learned about Sally's arrogance. This produces (and reinforces) the segregated status of privileged information in forming *social* evaluative estimates in particular. In contrast, when privileged information holders *themselves* evaluate a target, they have no reason to naturally (and simultaneously) consider two such evaluations. As such, the segregated processing account sees the PIE as a distinctly *social* phenomenon: Holding

privileged (vs. non-privileged) information should not tilt the self's own evaluations. But by the alternative pure focalism account, privileged information influences judgments merely because it is salient or top-of-mind, and is thus especially influential on any judgment. By this alternative hypothesis, Josh's *own* evaluation of Sally would also be more negative when knowledge of her arrogance is privileged instead of common knowledge. After all, previous research explained by focalism mechanisms typically demonstrate how salient information can color one's *own* judgments of a target (O'Brien, Kristal, Ellsworth, & Schwarz, 2018; Schwarz, Kahneman & Xu, 2009).

Second, the segregated processing account identifies a boundary on when the PIE should emerge because of *how* privileged information is incorporated into SEEs. That is, when considering the potential closing of an information asymmetry, a privileged information holder separately incorporates the privileged information into a preexisting SEE. Contrast this with the mirror process of counterfactually opening an information asymmetry (e.g., "How would Colleen evaluate Sally if she had *not* learned about Sally's arrogance?"). In this case, the (counterfactually) privileged information was already processed and incorporated into the SEE on equal footing with the other information. As a result, opening (instead of closing) an information asymmetry should not lead this focal information to disproportionately sway SEEs. In contrast, the pure focalism logic—which sees the PIE emerging merely because one changes a SEE in light of that (focal) attribute in particular—would predict that counterfactually opening information asymmetries would produce an effect parallel to the PIE. By this alternative focalism hypothesis, if Josh were to imagine that Colleen *hadn't* already learned about Sally's arrogance, then he would imagine that Colleen's evaluation of Sally would be even more positive than what

he assumed the third committee member (who has never learned this information) currently thinks of Sally.

A third prediction suggests a more direct test of the segregated processing account. Specifically, even though holding privileged information naturally encourages it to be incorporated into an SEE in a segregated manner, external intervention could encourage a less segregated (and thus, more integrated) incorporation. This general methodological approach complements previously used interventions that modify judges' natural processing approach to establish the importance of specific information-processing mechanisms (e.g., Cooney, Boothby, & Lee, 2021; Moon, Gan, & Critcher, 2020; Savitsky, Epley & Gilovich, 2001). In this same vein, we will have some privileged information holders process the privileged information *in light of* (and thus as one among the other pieces of) information that is common knowledge. If the segregated processing account helps to produce the PIE, then interfering with this segregated processing (by encouraging integration) should reduce the PIE.

Where this third prediction homes in on the conditions which give rise to the PIE, our fourth prediction pinpoints exactly *what* produces the PIE. Here, we address a different alternative account: Research in the secrecy literature might anticipate an *attribute extremity* mechanism is what produces the PIE. More specifically, keeping information hidden can lead people to extremify that information's valence (see Lane & Wegner, 1995). Thus, information that only the self knows—positive or negative—might feel more evaluatively extreme than information that is common knowledge. In contrast, we argue that an *enhanced weighting* mechanism produces the PIE. That is, when information is processed in a segregated manner, its influence should not be watered down by being incorporated as one attribute among many, meaning it should be weighted more heavily in the formation of an SEE (than if that same

information were non-privileged). Of course, both of these possibilities—enhanced weighting and evaluative extremity—are not mutually exclusive, so we will test both directly.

Overview of Studies

Across seven studies, we test for a *privileged information effect* (PIE): Social evaluative estimates will be more strongly influenced by information—positive or negative—that people consider *adding* to their understanding of what another knows (e.g., because it is currently privileged) compared to if that information has been believed to be known by another all along. In our basic paradigm, participants make SEEs about another person (often said to be another participant)—one who did not yet have access to privileged information held by the self, or one who already had the same information as did participants themselves. Studies 1 and 2 tested for the PIE using ecologically valid materials. In Study 1, participants saw actual video ads that had or had not been edited to make certain information privileged. In Study 2, participants tried a new technology product (i.e., a virtual reality headset) in the lab before writing a to-be-shared product review that did or did not mention a subtle product flaw. In both studies, we predicted that SEEs would be guided more by information that was initially privileged (as opposed to common knowledge). Study 3 then tested the extent to which the privileged information effect is indeed a social bias (i.e., whether initially withheld information actually does exert as strong of an effect on someone’s evaluations as privileged information holders expected).

Studies 4-6 were designed to test whether the segregated processing account, as opposed to the focalism account, does indeed predict the nature and scope of the PIE. Study 4 tested whether privileged information influences not merely social evaluative estimates but also the self’s own evaluations. Study 5 unconfounded mentally *closing* an information asymmetry from *opening* one. Study 6 most directly tested the segregated processing account by introducing a

new condition that preserved privileged information's privileged status but encouraged participants to process and incorporate the privileged information into their SEEs in a non-segregated (i.e., integrated) manner.

Finally, Study 7 returned to a more naturalistic context in which to test the PIE, while also testing a mechanistic implication of our reasoning. More specifically, participants composed matchmaking letters to a single friend or acquaintance in which they included or withheld information about a dating prospect. We tested whether the PIE would emerge because: (1) privileged information—more so than common knowledge—was *weighted* more heavily in privileged information holders' SEEs, and/or (2) the valence of privileged information would itself start to seem more extreme. In this way, Study 7 tests *how* privileged information exerts an undue influence on SEEs.

Consistent with Simmons, Nelson, and Simonsohn (2013), we report how we determined our sample size, as well as all data exclusions (if any), manipulations, and measures. For all studies, we predetermined a stopping rule for data collection. For each of the laboratory studies, we collected as many participants as we could in one or two academic semesters. For the online survey studies, we maximized our sample size in light of the lab's funding resources (distributed across all studies expected to be run that month.) Mindful that researchers typically do not know true effect sizes in advance of data collection, Simmons, Nelson, and Simonsohn (2013) endorse the creation of *ex ante* stopping rules that—without strong justification otherwise—allowed researchers to achieve at least fifty participants per cell. On average, our studies achieve sample sizes of 170 per cell. Data, materials, and analysis scripts for all studies can be found here: https://osf.io/t7dfp/?view_only=d7ba950c8cd6495bb929d1948e6eb90e.

Study 1

Study 1 offered an initial test of the privileged information effect. Participants in Study 1 evaluated Cabin, a little-known travel company headquartered in California's Bay Area.

Although its services were suspended at the start of the COVID-19 pandemic, Cabin offered luxury transportation between San Francisco and Los Angeles. The imagery on Cabin's website is meant to reflect a slick, posh, upscale experience. It is then perhaps not surprising that Cabin seems to go out of its way to not reveal that they are, in fact, a bus service. Clearly, Cabin does not want to be seen as a substitute for other bus companies (e.g., Greyhound) that offer cheaper and lower-quality travel experiences.

Participants in Study 1 were ostensibly yoked to another participant who also learned about Cabin. All participants saw an actual complete video advertisement from Cabin that did (briefly) include imagery making clear that Cabin is indeed a bus. But we varied whether the yoked other also supposedly saw either the complete ad or a shortened version of the ad in which this (presumably negative) information had been edited out. This made knowledge that Cabin was indeed a bus service *privileged* (i.e., known only to the participant) or *non-privileged* (i.e., known to the yoked other as well). Participants then estimated how the yoked other—either based on their already complete information (non-privileged condition) or once they saw the full-length version of the ad (privileged condition)—would evaluate Cabin. In this way, all participants provided SEEs about another whom they assumed had the same, complete information as the self. We hypothesized that privileged information holders would estimate that the other would judge Cabin more negatively compared to those who thought this (negative) information was already common knowledge. If so, this would provide an initial demonstration of the PIE.

Method

Participants and design. Four hundred one participants took part in exchange for course credit or payment. In an effort to achieve a large sample size, we recruited participants simultaneously from a subject pool at an American university as well as Amazon Mechanical Turk (AMT). Participants were randomly assigned to one of two information conditions: *privileged* or *non-privileged*. Fifty-one participants failed an attention check (see Supplemental Materials) and—following our preregistration—were excluded from all analyses reported below. This left a final sample of 350 for all analyses. The preregistration—which includes details on this exclusion criterion, as well as the methods, sample size, and hypothesis—can be found here: https://osf.io/t7dfp/?view_only=d7ba950c8cd6495bb929d1948e6eb90e. Information about an additional variable, which we preregistered as exploratory, can also be found in the Supplemental Materials.

Materials and procedure. Participants learned they would be assisting a company called Cabin. We assumed most participants were not already familiar with this startup, so participants were told, “Cabin is a travel company that offers medium-haul travel between major U.S. cities (e.g., San Francisco and Los Angeles).” In fact, at the time the study was run, Cabin was only operating between those two cities. To provide a high-level overview, participants were given the following information about their task:

“Cabin is currently testing different advertisements that would play before videos on online platforms such as YouTube and Hulu. Because these platforms have space for advertisements of varying lengths, Cabin is currently testing shorter and longer versions of their advertisements. Cabin likes to test these advertisements on various groups of people in the hopes of crowdsourcing the evaluation of which advertisements will be successful.”

Next, participants learned that—given Cabin’s supposed interest in testing how “presenting information in different ways may change evaluations of their company”—each participant in the study would see one of two versions of an ad: a shorter or a longer version. Of course, our true goal was not to evaluate participants’ *own* responses to the two ads, but to devise a cover story to explain why participants would judge how another would respond to information that was or was not currently privileged for the self. Toward this end, we led participants to believe they were being yoked to another participant who had just arrived at the task. (In actuality, there was no such participant.) A loading-screen animated gif indicated that the computer was randomly selecting which ad each participant—the self and the yoked other—would see.

All participants learned they had been randomly assigned to watch the longer version of the advertisement. Which ad the yoked other would supposedly see varied by information condition. When the self was to have privileged information, the yoked other would watch the shorter version. When the self was to have non-privileged information, the yoked other would see the longer version.

At that point, all participants watched the longer (28 s) version of the Cabin ad. This video ad was pulled from Cabin’s website. The video is a sleek marketing production that shows luxurious amenities, spacious lounge areas, a chic modern bathroom, and futuristic sleeping cabins. But what one might not know—were it not for imagery shown in the first three seconds of the ad—is that Cabin is a bus service. Given the aesthetic and décor of the experience seem to fit more with what one might expect to see on a private luxury jet, we suspected these initial three seconds offered something of a contaminating feature. And indeed, an inspection of

Cabin's website suggests they have the same hunch: They certainly do not lie that they are not a bus service, but they seem to go to pains not to emphasize this fact.

At this point, participants' experience varied by their information condition. Those in the privileged information condition saw two different videos—the shorter version (25 s) that would be shown to the other participant, as well as the 3 s of content that had been edited out to create the shorter version. Those three seconds—which featured a zoomed-out view of a Cabin vehicle—were the only hints that Cabin was indeed a bus service. To make certain that participants understood which content the yoked participant would (and would not) see, we asked participants to summarize the content of each video. To hold constant participants' exposure to the total content of the ad, those in the non-privileged information rewatched the longer version of the ad and also summarized its content.

All participants then completed two measures that assessed their estimates about how the yoked other would evaluate Cabin (i.e., their SEEs). But crucially, these measures varied slightly by condition. For those in the non-privileged information condition (those whose yoked other saw the full version of the ad), they were merely asked to guess how the yoked other would evaluate Cabin and how interested the yoked other would be in using Cabin's services in the future. These responses were made on 101-point slider scales anchored at -50 (*very negatively / extremely uninterested*) and +50 (*very positively / extremely interested*). Each scale defaulted to 0.

Privileged information participants completed these same two measures, but under slightly different instructions. They were asked to “imagine that the other participant were to also see” the information that was edited out of the short ad. We then added the words “If so:” before each measure, so privileged information participants understood they should make these SEEs

under the potential reality in which the yoked other had full information. In other words, and crucially, participants in both conditions made SEEs about a yoked other who had the same information about Cabin. But for those in the privileged information condition, some of that information (in particular, the negative detail) was currently hidden from the yoked other. Across both conditions, these two items were averaged ($r = .66$) to form an *evaluation* composite.

Results and Discussion

We wanted to test whether participants holding privileged (negative) information assumed that the yoked other—if they were to have that same information—would evaluate Cabin more negatively than would participants for whom this information was non-privileged. Consistent with our preregistered hypothesis, participants in the privileged information condition believed that the yoked other would evaluate Cabin less positively ($M = 16.50$, $SD = 23.50$) than participants in the non-privileged information condition ($M = 28.47$, $SD = 14.52$), $t(296.56) = 5.75$, $p < .001$, $d = 0.61$ (see Table 1). This provides initial evidence for the PIE: Information about Cabin being a bus service was assumed to be more influential in others' evaluations when it had yet to be disclosed to them (as opposed to being known by them all along).

It was important in Study 1 that all participants made SEEs about someone who knew (or would know) the same information about Cabin that was shown to the self. Critically,

Table 1

SEEs of Yoked Others' Evaluations of Cabin, by Information Condition (Study 1)

SEE	Information Condition	
	Non-privileged	Privileged
Evaluation	32.17 (13.92) _a	17.84 (25.54) _b
Interest	24.76 (17.73) _a	15.16 (26.61) _b
<i>Evaluation Composite</i>	28.47 (14.52) _a	16.50 (23.50) _b

Note. Means (with accompanying standard deviations) in the same row that do not share the same subscript differ at the $p < .01$ level.

participants were exposed to all of this information twice, regardless of their condition. That said, privileged information was actually segregated in two ways. First, privileged information was *psychologically* segregated (as it always is): It differentiated what another currently knew from what another might ultimately know. But second, the information was *presented* in a uniquely segregated form: Privileged information holders were unique in seeing the three-second (negative) portion of the video in a stand-alone format. In the remaining studies, we will implement the privileged information manipulation in a way such that all participants consider the (sometimes) privileged information as a discrete informational chunk, even though only some participants would have reason to segregate the information in the formation of their SEE.

Study 2

Study 2 moved to a new context in which asymmetric information might arise. Those who have actually used a product necessarily have more information about it than those who are merely shopping for one. After all, that is one reason why shoppers read reviews. In this vein, we had all participants in Study 2 test out a new technology product, the Oculus Go.

The Oculus Go was released in 2018 as part of the “first wave” of Facebook Technologies’ VR devices. The headset was marketed to consumers as a cost-friendly streaming and gaming device; initial advertisements for the device promised an immersive virtual reality experience in which users could “go virtually everywhere.” Although the headset’s crisp video graphics and extensive array of available content allow users to inhabit realistic virtual worlds, the device has a subtle (and mildly disruptive) flaw. After users put the headset on and direct their eyes toward the base of the contraption, they can see a sliver of light that creeps through the bottom of the device. This renders the VR experience slightly less immersive than advertised.

Following a trial experience with the Oculus Go, participants wrote a structured review of the product that would ostensibly be shared—in full or in part—with another participant. In this way, participants were not selectively exposed to information that was sometimes withheld from another (as in Study 1), but instead authored information themselves that was disclosed fully or only in part to another. In this way, we test for the PIE's robustness by determining whether it emerges even when it is the self (instead of a third party) who is supplying complete or incomplete information to another.

That said, this paradigm loses a bit in terms of standardization, for each participant wrote their own reviews. To constrain (theoretically irrelevant) variability, we led participants through a structured, review-writing process. More specifically, participants were asked to review the headset by elaborating on a number of its features: three positive and one negative. In this way, all participants considered as an individual piece of information the feature that was sometimes withheld (i.e., the negative feature). All participants were ostensibly yoked to a future participant who would receive some or all of the participant's review and then offer an evaluation of the Go. In the *non-privileged* information condition, participants' entire review would be shared. In the *privileged* information condition, participants' elaboration on the Go's negative feature would be withheld. Then, participants estimated how the yoked other would evaluate the Oculus Go with all of the information from the review (non-privileged) or *if they were to have* all of the information from the review (privileged). We predicted that SEEs of the Oculus would be more negative when the negative product feature was first withheld from (privileged information condition) as opposed to immediately shared with the other (non-privileged information condition). This pattern would support the robustness and generalizability of the privileged information effect.

Method

Participants and design. Three hundred forty undergraduates at an American university were randomly assigned to one of two information conditions: *privileged* or *non-privileged*. Seventy-five participants failed an attention check (see Supplemental Materials) and—following our preregistration—were excluded from all analyses reported below. This left a final sample of 265 for all analyses. The preregistration can be found here:

https://osf.io/t7dfp/?view_only=d7ba950c8cd6495bb929d1948e6eb90e.²

Materials and procedure. Upon arriving at the laboratory, participants were led to private rooms where they were seated in front of a laptop. There was an Oculus Go virtual reality headset and handheld remote on the table in front of them. All participants were told that they would be assisting a company called Laurel Technologies. Participants read that: “Laurel Technologies is a company that evaluates the effectiveness of advertising strategies. Specifically, Laurel Technologies evaluates how consumers' reviews of products can be used in combination with reviews from prominent media outlets to advertise products.” Then, participants were given the following information about the task that they would be completing:

“Laurel Technologies is currently interested in how trial experiences influence word-of-mouth communication between consumers, as well as how presenting consumers with different information influences how consumers evaluate products. Laurel Technologies is interested in consumers' evaluations of the Oculus Go—a new virtual reality headset. In this session, you are in a condition in which you are a 'reviewer.' Reviewers will experience the Oculus Go (on

² We initially preregistered our intent to recruit as many participants as we could during one semester. Unfortunately, the instructor miscommunicated to students partway through the semester the nature of the subject pool requirement, which undermined recruitment. As a result, we extended data collection an additional academic term.

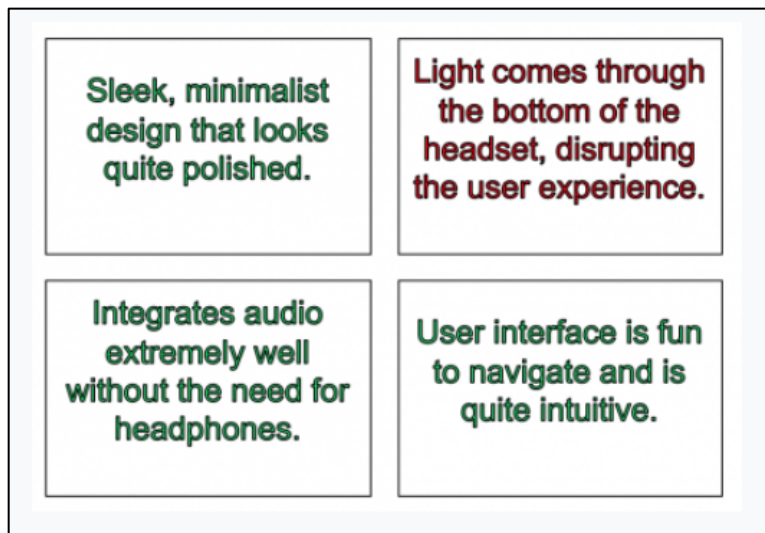
the table) for a brief trial period before making comments that will be sent to a participant in the next session of this research study (the other participant).”

All participants were thus told that they would be: (1) using the Oculus Go for a brief trial experience, and (2) making comments about their trial experience that would ostensibly be sent to a future participant.

To assist with the creation of their own review, we showed all participants four rectangular cards (see Figure 1). Each card contained information about a different feature or aspect of the Oculus Go. After reading each of the cards, participants were instructed to examine the Oculus Go in order to consider each of the points in more detail. Three of the cards were

Figure 1

Oculus Go Features Displayed to Participants (Study 2)



Note. Participants in Study 2 saw four cards, each of which corresponded with one of the Oculus Go’s features. Three of the cards (with green text) highlighted positive features, while one of the cards (with red text) highlighted a negative feature. Each feature was itself adapted from actual professional product reviews about the Oculus Go.

green and detailed features that were positive—that is, enjoyable or exciting features that users might mention in a favorable review of the Oculus Go. The other card was red and detailed a less positive feature of the Oculus Go, that light comes up through the bottom of the headset, thereby disrupting the user experience. When considering this card, participants were instructed to put the headset on and look down in order to notice how light seeps through the bottom of the headset.

Next, all participants began their 5-minute trial experience with the Oculus Go. After reading a brief description of how to use the device’s handheld remote, participants listened to an audio recording that guided them through the trial. Given participants could not see the computer while wearing the headset, the audio presentation allowed participants to simultaneously listen to the instructions and use the Oculus Go. As part of the trial experience, participants navigated through the Oculus Go’s home screen and played a game called Smash Hits, in which players move through a multicolored passageway and shoot metal marbles at various obstacles and targets.

After completing the trial experience and removing the headset, all participants again considered the four cards that detailed four features of the Oculus Go. For each feature, participants were asked to draw on their trial experience and comment on how that feature was positive or enjoyable (for the three green cards) or negative and bothersome (for the single red card). Participants wrote their comments in four different textboxes — each of which corresponded with one of the four features.

All participants then read that they would see which of their comments would be shared with the other participant. Whether some or all of the comments would (ostensibly) be passed along varied by information condition. Participants in the *non-privileged* condition read that the

yoked other would see each of the four feature-cards as well as participants' corresponding comments about that feature. Participants in the *privileged* condition instead learned that the yoked other would not see the fourth and final card—the one that conceded that light comes through the bottom of the headset—or participants' comments about that feature.

Participants then completed a measure that asked them to estimate how the yoked other would evaluate the Oculus Go. As in Study 1, this measure varied slightly by condition. Participants in the non-privileged information condition—for whom the yoked other ostensibly had complete information—estimated how the yoked other would evaluate the Oculus Go on a 101-point slider scale anchored at -50 (*very negatively*) and +50 (*very positively*). Participants in the privileged information condition were asked to “imagine that the other participant were to also see” the fourth card (about light creeping in through the bottom of the Oculus headset) and participants' comments about it. As in Study 1, we then added the words “If so:” before the dependent measure; thus, participants in the privileged condition were making SEEs that mirrored participants' SEEs in the non-privileged information condition. That is, participants in both conditions made SEEs about a yoked other who was assumed to have the same information about the Oculus Go.

Results and Discussion

When knowledge of the product's flaw was privileged, did participants assume that it would lead to more negative product evaluations than when that information was common knowledge? In short, yes: We again found evidence of a PIE. Participants in the privileged information condition believed that the other participant would evaluate the Oculus Go less positively ($M = 25.98$, $SD = 19.37$) than did participants in the non-privileged information condition ($M = 31.44$, $SD = 11.44$), $t(223.16) = 2.82$, $p = .005$, $d = 0.34$.

One notable feature about Study 2 is that—unlike in Study 1—participants were not led to believe that the other participant was participating at the exact same time as they themselves were. In that sense, all participants in Study 2 considered a person who would learn information about the Oculus Go in the future. This reinforces that what is important in producing the PIE is that people are considering a target who has already learned or who has yet to learn a key piece of information, not the literal timing of when this asymmetry is resolved. Thus, taken together, Studies 1 and 2 offer evidence of the PIE when participants estimate the evaluations of naturalistically rich stimuli. We next turn to four studies that—via more tightly controlled paradigms—permit us to better understand the nature and boundary conditions of the PIE.

Study 3

Study 3 builds on our previous studies in three primary ways. First, in contrast with Studies 1 and 2, the privileged information was positive. One alternative interpretation of Studies 1-2 is that the results do not reflect the greater assumed influence of privileged information, but instead heightened negativity that might be assumed to stem from having received incomplete information. After all, feeling that a secret is being kept from the self can be grating (Jones, Carter-Sowell, Kelly, & Williams, 2009). This force might lead privileged information holders to assume that people will sour on a target upon learning (previously) privileged information. Conversely, if, as we propose, the PIE reflects the assumed disproportionate influence of privileged information, then privileged information holders' SEEs in Study 3 should skew positive. Second, we wanted to address a concern that there might be some signal value in what piece of information was kept privileged. For example, withholding information about the Oculus's flaw may have implicitly flagged the problem as especially severe (see Travers, Van

Boven, & Judd, 2014). In the next three studies, we use a paradigm in which participants are told that which of four pieces of information was (sometimes) privileged was randomly determined.

Third, we wanted to understand whether the PIE is indeed a social bias. That is, does initially withheld information actually exert as sizeable of an effect on another's evaluations as privileged information holders expect? Thus far, we have established the PIE by comparing SEEs about another person who has the same (complete) information about a target. What has varied is whether the other is considered as *currently* or *potentially* having this complete information. Note that, in the privileged information condition, the other—unlike the self—is imagined to learn the information in two stages: common knowledge is learned initially, and privileged information is then (imagined to be) subsequently acquired.

This raises the possibility that the PIE may not reflect a normative error in social judgment, but instead an accurate lay theory about how others incorporate (previously privileged) information. Although such a phenomenon would still be intriguing in its own right, Study 3 addressed this possibility directly by testing whether privileged information holders actually do overestimate the influence of information that is initially withheld from another. We accomplish this by assigning some participants to the role of the person from whom privileged information is initially withheld. We can then see to what extent—when these people ultimately do acquire full information—that: 1) this information does exert a larger influence on their evaluations (than it would have had it been learned along with the rest), and 2) whether any such influence is as large as privileged information holders (as reflected in their SEEs) expect it to be.

Method

Participants and design. Nine hundred fifteen Americans were recruited from AMT in exchange for nominal compensation. Participants were randomly assigned to one of four

conditions in a 2(information: privileged or non-privileged) X 2(perspective: self or social) full-factorial design. Eighty-two participants failed a memory-based attention check (see Supplemental Materials) and were excluded from all analyses reported below. This left a final sample of 833 participants.

Materials and procedure. All participants first read that they would be assisting a venture capital company (Heritage Capital) interested in forecasting stock performance. To that end, participants then learned that they would be evaluating an investment opportunity. All participants received the following information:

“Heritage Capital wants two separate people to evaluate different features of each investment opportunity. This means that you will be matched with the next MTurker who accepts this HIT and both of you will be evaluating: Cedar Technologies Inc. Heritage Capital is curious about how presenting information in different ways may change people’s evaluation of the investment opportunities.”

Then, participants learned that both they and the yoked other would receive information about different features of the investment opportunity. All participants saw four cards (see Figure 2). Each card possessed a(n initially hidden-from-view) feature about the investment opportunity. We conducted a pretest (N = 106 Americans on AMT; full details reported in Supplemental Materials) to identify two features that were relatively neutral, one that was relatively positive, and one that was relatively negative. In the pretest, we had participants evaluate twenty-one features, as we were in search of a positive and negative attribute that were roughly equidistant from the evaluatively neutral midpoint.

Information manipulation. Depending on participants’ information condition, they either learned that both they and the yoked other would learn all four pieces of information about the

Figure 2*Investment Opportunity Features Displayed to Participants (Study 3)*

Feature A	Feature B	Feature C	Feature D
Several analysts have encouraged current owners of the stock to "hold" the stock – that they should neither buy nor sell the stock.	The company has a massive manufacturing plant located in China.	There have been rumors that Google may be interested in acquiring the company.	A mid-level executive at the company just quit following a sexual harassment lawsuit.

Note. Participants in Study 3 saw four rectangular cards; each card corresponded with one of the investment opportunity's features. Feature C, which was always presented last, was sometimes said not to be shown to the other person.

investment opportunity (*non-privileged* information condition) or that one person would receive three pieces of information while the other received all four (*privileged* information condition). To begin, all participants saw three pieces of information in a random order: the two neutral and one negative feature. At that point, all non-privileged information participants learned the fourth and final (positive) feature ("There have been rumors that Google may be interested in acquiring the company"). Exactly *when* participants in the privileged information condition learned this final piece of information depended on the perspective manipulation. This is because, as described in more detail below, the perspective manipulation determined who (the participant or the yoked other) was initially blind to the privileged feature.

Perspective manipulation. The perspective manipulation determined whether participants would offer their own evaluations of the investment opportunity (*self* perspective condition) or

make SEEs about the yoked other (*social* perspective condition). This was straightforward in the non-privileged condition: Participants offered their own evaluations or estimated the evaluations of a yoked other who also possessed all information (i.e., all four features) about the target.

In the privileged information condition, the perspective manipulation also determined the nature of the initial information asymmetry. Privileged information participants in the self perspective condition were initially kept in the dark about the fourth feature. That is, they knew that the feature was shown to the yoked other, but that it was hidden from the self. For privileged information participants who would estimate the yoked other's evaluation, this final feature was shown to the self while ostensibly withheld from the other. In other words, privileged information participants who would offer the self's or estimate the yoked other's perspective were initially blind to or uniquely in possession of knowledge about the fourth feature, respectively. Next, the information asymmetry was actually resolved (in the self perspective condition) or, like in our previous studies and for the purpose of evaluating the privileged information effect, it was (in the social perspective condition) hypothetically resolved ("Imagine that the other person did learn about the [privileged feature]..."; social perspective condition).

The evaluation composite. Regardless of whether participants offered their own evaluation of the investment opportunity or estimated the evaluation of the yoked other, participants completed two items. One asked how participants themselves—or the yoked other—would evaluate the investment opportunity overall. This was made on a 101-point slider scale anchored at -50 (*very negatively*) and +50 (*very positively*). The second item asked participants to indicate how likely participants themselves—or the yoked other—would be to invest their own money in Cedar Technologies, Inc. This response was given on a 101-point slider scale

anchored at 0% and 100%. After placing these items on equivalent, 0-to-100 scales, these items were averaged to form an *evaluation* composite ($r = .76$).

We note one methodological difference between the non-privileged and privileged information conditions. Those in the non-privileged condition completed these items once, after everyone had learned all four features of the target investment opportunity. Those in the privileged condition instead completed these items twice: once during the information asymmetry, and (crucially) then again after this asymmetry was resolved. Given the former measures were collected on or about people who did not yet hold all of the relevant information, it is only the latter measurement that factors into our analyses. In this way, all analyses were on the evaluation composite that was about someone who actually had (or was assumed to have) all four pieces of information about the investment opportunity.

Results and Discussion

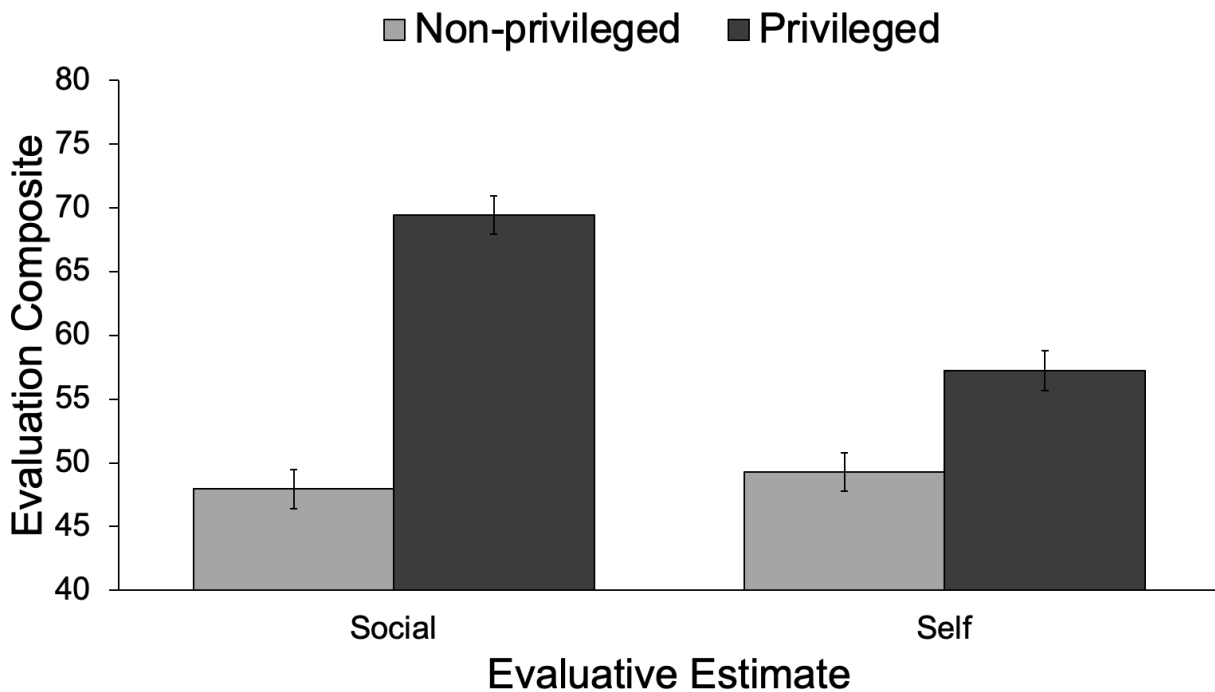
First, we examined the SEEs of those in the other perspective condition. Providing another replication of the PIE, participants estimated that others' evaluations would be more positive if they did not currently know the fourth feature (privileged information condition: $M = 69.41$, $SD = 18.55$) than if that information was never withheld from them (non-privileged information condition: $M = 47.94$, $SD = 22.76$), $t(421) = 10.67$, $p < .001$, $d = 1.03$. We proceeded to examine how much self perspective participants' own evaluations were influenced by having some information initially withheld with them. This would allow us to disentangle whether the PIE reflects: 1) an accurate lay theory, 2) a directionally accurate but exaggerated social bias, or 3) a completely misguided pattern of social inference (with no directional truth underlying it).

To begin to distinguish amongst these three possibilities, we submitted the evaluation composite to a two-way 2(information: privileged or non-privileged) X 2(perspective: self or

social) ANOVA. Suggesting that the PIE does not simply reflect an accurate SEE (possibility #1), a strong Information \times Perspective interaction emerged, $F(1, 829) = 19.45, p < .001, \eta_p^2 = .023$ (see Figure 3). But was there at least directional truth behind the PIE? There was (possibility #2): Participants in the privileged self perspective condition evaluated the investment

Figure 3

Participants' Evaluations or SEEs about the Investment Opportunity, by Information Condition (Study 3)



Note. Participants in Study 3 either made SEEs about the investment opportunity (social condition) or provided their own evaluations of the investment opportunity (self condition). These judgments were made either once the self and the yoked other had complete information about the investment opportunity (non-privileged information condition) or when only the self had complete information about the investment opportunity in a second stage (privileged information condition). Error bars represent ± 1 SEM.

opportunity more positively ($M = 57.22$, $SD = 24.34$) than participants in the non-privileged self perspective condition did ($M = 49.27$, $SD = 22.63$), $t(408) = 3.43$, $p = .001$, $d = 0.34$. In other words, when participants learned the fourth (and positive) feature of the investment opportunity in a second, segregated stage, their final evaluations were a bit more aligned with the valence of that once-withheld feature. But the significant interaction emerged because participants' SEEs overestimated this gap by more than a factor of two.

The interaction can be decomposed in a different way, which reinforces the same conclusion. When participants considered how another would respond to an investment opportunity characterized by positive information that had been initially withheld from them, they overestimated how positive such evaluations would be compared to how participants in that exact situation responded themselves, $t(419) = 5.81$, $p < .001$, $d = 0.56$. This reinforces that the PIE is a social error. But showing this was not a simple positivity bias in estimating others' evaluations, this self-other gap was eliminated when information was never privileged, $t < 1$.

In sum, Study 3 demonstrated that—although there may be some directional truth to the PIE—privileged information holders overestimated the impact of ultimately learning privileged information. This suggests that the PIE joins other social biases that may be based in kernels of truth that are then exaggerated or overly applied (Gigerenzer & Brighton, 2009; Moon, Gan, & Critcher, 2020; Yamagishi, Terai, Kiyonari, Mifune, & Kanazawa, 2007). However, having found consistent and robust evidence for the PIE, we sought to conduct an additional test to further explore the actual effect of once-privileged information on the newly informed's evaluations. Although we had hoped to more deeply probe the effect that suggested there may be some truth to the PIE, Supplemental Study A did not replicate the present study's finding that learning once-privileged information is particularly influential on one's own evaluations.

Tentatively, we think the PIE is fairly consistent and robust (as the subsequent studies will continue to reinforce), though once-privileged information's actual influence on people may be more fragile or at least variable. We wait until the General Discussion—armed with a more complete understanding of the nature of the PIE—to offer additional speculation on when and how privileged information may be most likely to influence the once-uninformed.

Study 4

Study 4 built on the previous studies in three ways. First, we moved to a new privileged information context. Participants learned that they would be helping a(n actually fictitious) company with the selection of a regional manager. All participants learned four attributes about one of the job candidates—three positive and one negative. To allow us to test for the PIE, this negative attribute, allegedly, would either be disclosed to the next participant (non-privileged information condition) or be initially withheld from them (privileged information condition). Second, in Studies 2 and 3, in order to reinforce our privileged information manipulation, we had such participants offer an SEE about another's *current* evaluation before considering what another's evaluation would be with complete information. Although we think that privileged information holders naturally form both types of SEEs (i.e., considering what another does know and what another could know), we wanted to affirm that this methodological feature is not a precondition for the PIE. For that reason—like Study 1—we omitted these initial SEE measures.

Third, Study 4 aimed to establish the scope of the PIE. Thus far, we have shown that privileged and non-privileged information holders make different SEEs of another's evaluations. What has not been examined is whether such findings are specific to these *social* evaluative estimates. We had some privileged information participants offer their own evaluation of the job candidate instead of estimating what another participant's evaluation would be. Non-privileged

information participants—those in what is essentially our control condition—also offered their own take on the job candidate. This allowed us to test whether a privileged information effect would characterize participants' own evaluations as well. Previous research has found that privileged information can easily become top-of-mind (e.g., Slepian, Greenaway, & Masicampo, 2020). If the PIE emerges merely because privileged information is focal, then a focalism account anticipates that privileged information will disproportionately influence not merely social evaluative estimates but the self's own evaluations as well. In contrast, the segregated processing account, which argues that the PIE emerges due to the segregated construction of another's representation of a target, anticipates that the PIE should apply only to SEEs.

Method

Participants and design. Three hundred thirty-nine Americans were recruited from AMT in exchange for nominal compensation. Participants were randomly assigned to one of three conditions: privileged/self-judgment, privileged/other-judgment, or non-privileged. Two conditions—privileged/self-judgment and privileged/other-judgment—involved participants' holding privileged information that another participant did not have. In one case, participants offered their own evaluation of the target (privileged/self-judgment); in the other, they predicted the yoked other's evaluation of the target (privileged/other judgment). Those in the non-privileged condition ultimately offered both their own evaluation (non-privileged/self-judgment) and estimated the yoked other's evaluation (non-privileged/other-judgment) of the target. Twenty-three participants failed an attention check (see Supplemental Materials) and were excluded from all analyses reported below. This left a final sample of 316 participants.

Materials and procedure. Participants learned that they would be completing a hiring exercise. In it, participants were asked to assist a company (Heritage Restaurants) that was hiring a new regional manager. Participants received the following information about the exercise:

“An independent recruiting agency has offered up a number of candidates as potential hires for the position. Heritage Restaurants wants two separate people to evaluate a candidate before corporate headquarters reviews the top candidates. This means that you and the next MTurker who accepts this HIT will be evaluating: Michael Clark.”

Participants read that both they themselves and the yoked other would receive multiple pieces of information about Michael Clark (the job candidate). Those in the *non-privileged* condition read, “Both you and the other person evaluating this candidate will receive 4 pieces of information about the candidate.” Conversely, participants in the two *privileged* conditions read, “One of you—you or the other person evaluating this candidate—will receive 4 pieces of information about the candidate. The other of you will receive 3 pieces of information about the candidate.” In this way, participants always saw all 4 pieces of information about the job candidate. What varied was the amount of information that the yoked other was said to receive.

Next, all participants saw four rectangular cards. The back of each card contained one of the four pieces of information about the job candidate. One by one, each of the cards was (virtually) flipped to reveal a distinct attribute. As each feature was presented, participants learned whether the yoked other would also be shown that card. In the non-privileged condition, participants learned the yoked other read each card as well. In the two privileged conditions, participants learned that only three of the attributes would be common knowledge: Those spoke to the job candidate’s emotional intelligence, talent at public speaking, and efficiency in email communication. When the one (negative) privileged attribute (i.e., Attribute C: The job

candidate occasionally loses his temper and yells at subordinates) was presented, participants read, “The following attribute...will NOT also be shown to the other participant.”

Participants—regardless of whether they were offering their own opinion of the target or estimating the yoked other’s—next completed two measures. One asked for an evaluation (the self’s own or an SEE) of the job candidate. This judgment was provided on a 101-point slider scale anchored at -50 (*very negatively*) and +50 (*very positively*). The second asked how likely the self or the yoked other would be to offer the candidate a job with Heritage Restaurants. Participants responded on a 101-point slider scale anchored at 0% and 100%.

Participants in the non-privileged information condition answered these questions twice. First, they provided their own evaluations of the candidate (self-judgments). Next, they estimated the yoked other’s evaluations (other-judgments). Participants in the privileged/self-judgment condition answered only the self-judgments. In this way, they answered the same (initial) questions as those in the non-privileged condition, but did so with the understanding that they themselves possessed privileged (negative) information about the target. Participants in the privileged/other-judgment condition instead estimated the yoked other’s evaluation of the candidate. But like in our previous studies, these SEEs were made assuming the yoked other *were* to have complete information about the target—i.e., if they also knew the privileged attribute. The two items were tightly correlated, $r = .71$. After transforming both items to be on 0-to-100 scales (such that higher numbers indicate more positive actual or estimated evaluations), we averaged them to form an *evaluation* composite.

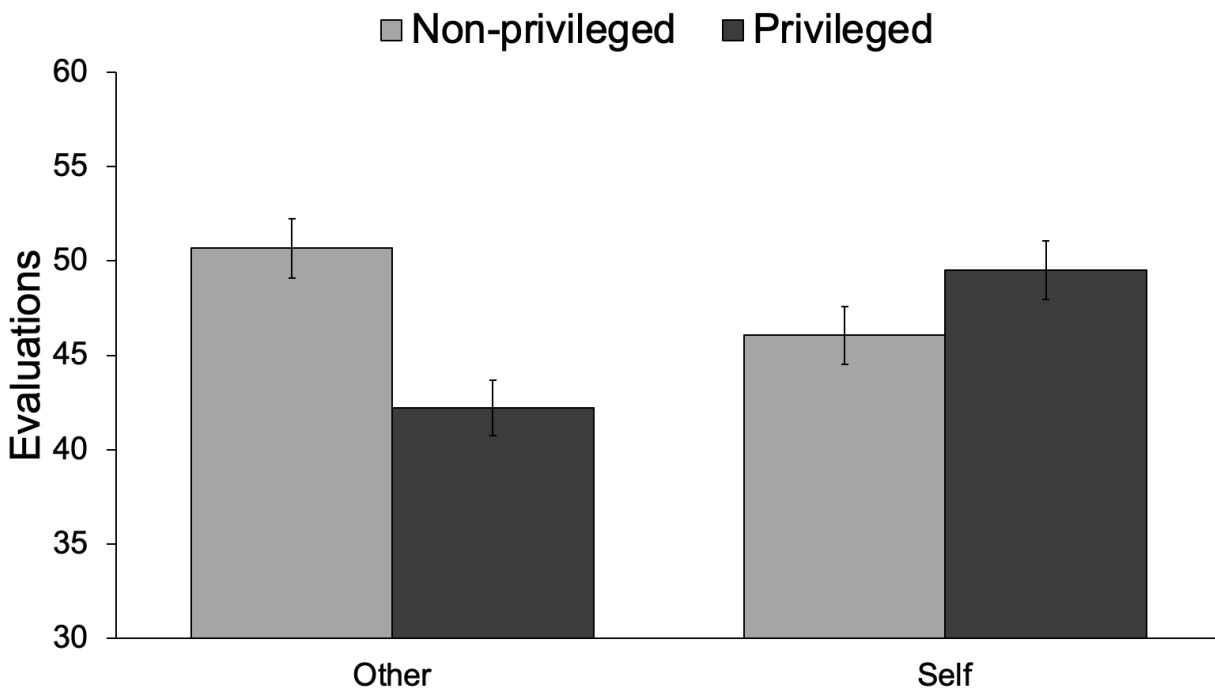
Results and Discussion

Participants in the privileged/other-judgment condition thought that the yoked other would evaluate the job candidate more negatively ($M = 42.22$, $SD = 22.24$) than did participants

in the non-privileged information condition ($M = 50.67$, $SD = 23.96$), $t(220) = 2.70$, $p = .007$, $d = 0.37$ (see Figure 4). This replicates—in a new domain—the PIE observed in Studies 1-3. Next, we considered whether participants themselves offered more negative evaluations of the target when they held (negative) privileged information—as opposed to information that was common knowledge—about the target. This was not the case: Participants’ self-judgments were not more negative (and were actually non-significantly more positive) when they held privileged information ($M = 49.51$, $SD = 23.85$) compared to when such information was non-privileged ($M = 46.05$, $SD = 25.86$), $t(214) = 1.10$, $p = .315$, $d = 0.14$. In other words, holding privileged

Figure 4

Participants’ Evaluations or SEEs about the Job Candidate, by Information Condition (Study 4)



Note. The left half of the figure depicts the two means relevant to testing for the PIE. The right half of the figure depicts the two means relevant for testing whether privileged information is focal in the self’s own evaluations of a target. Error bars represent +/- 1 SEM.

information changed how participants estimated others' SEEs, but not how they themselves evaluated the target. This reinforces that the PIE is a social phenomenon that stems from people's attempt to mentally close another's information deficit, not a phenomenon that merely reflects privileged information becoming unconditionally focal in any judgment that a privileged information holder makes.

Study 5

Study 5 built on our previous studies in three ways. First, we moved to a new judgment context: homebuying. Participants always received four pieces of information about a home: one relatively positive, two fairly neutral, and one relatively negative. Some (but not all) participants were led to believe that they, but not a yoked other, were solely aware of one of these four attributes: either the positive or the negative one. Second, although our previous studies suggested that the PIE emerges in contexts in which the privileged attribute is negative (Studies 1-2, 4) or positive (Study 3), Study 5 allowed us to test for both effects in the same judgment context. This would address a worry that privileged information—independent of its valence—might simply influence SEEs in a positive or negative direction in different judgment contexts.

Third, and reflective of our primary goal, Study 5 built on Study 4's experimental approach by taking another approach to empirically distinguish support for the segregated processing account from a mere focalism account. The present study considers whether it is closing an information asymmetry by *adding* information to an SEE—as opposed to merely manipulating an SEE in light of a focal piece of information—that leads the information to strongly guide SEEs. The segregated processing account argues that closing information asymmetries should be special. That is, when information is privileged, it is not incorporated into an SEE as one attribute among many. Instead, it is segregated from the common knowledge and

then incorporated on unequal (and thus privileged) footing. The pure focalism account would instead argue that the determinative factor is that this privileged information is focal, a component of an SEE whose (potential) influence is considered in isolation. The present study takes advantage of the fact that these two features are not necessarily confounded.

Study 5 thus introduces a subtle twist on our basic paradigm. For some participants, we again asked them to *close* an information asymmetry. But for other participants, we asked them to *open* an information asymmetry. That is, instead of asking, “How will this person feel about X when they learn Y?”, we instead asked, “How would this person feel about X had they not learned Y?” In that way, Y is focal in both cases, but only in the latter *subtractive* case was Y incorporated into the SEE on equal footing with the other features, as one attribute among many. Here, the segregated processing account does not anticipate that mentally subtracting it would be especially influential. In other words, these subtractive SEEs should not be all that dissimilar from SEEs about someone else who never learned that (counterfactually privileged) information to begin with. In contrast, the pure focalism account would not distinguish between closing an information asymmetry versus opening one. Both tasks require one to focus on a piece of information, and this focal status may lead it to be especially influential.

Method

Participants and design. Four hundred ninety-five Americans were recruited from AMT in exchange for nominal compensation. Participants were randomly assigned to one of four conditions in a 2(information: privileged or non-privileged) X 2(key feature: positive or negative) full-factorial design. Thirty participants failed an attention check (see Supplemental Materials). The remaining 465 participants are included in all analyses below.

Materials and procedure. To begin, participants learned that they would be assisting with the development of an app that could be used by prospective homebuyers. The app developers were supposedly interested in crowdsourcing evaluations of different homes. At that point, participants received more specific information about their own role in this process:

“Heritage Technologies wants two separate people to evaluate different features of each house. This means that you and the next MTurker who accepts this HIT will be evaluating: 1571 Cedar Street. Heritage Technologies is curious about how presenting information in different ways may change people’s evaluations of the houses.”

Next, participants read that both they themselves and the yoked other would receive information about several features of 1571 Cedar Street. To begin, all participants saw four cards, each of which was said to possess an attribute about the house on the other (currently not visible) side. A pretest ($N = 77$ American AMT workers; complete details in the Supplemental Materials) was used to identify four appropriate features. Pretest participants evaluated twenty-one features; we selected two that were relatively neutral (“The house has a very nice oak tree in front of it.”; “The living room has an area that would be perfect for two large bookshelves.”), one that was positive (“Solar panels on the roof keep the electricity bill low.”), and one that was negative (“The kitchen has a shortage of cabinet and storage space.”). The positive and negative features were roughly equidistant from the two neutral features. Participants in the *non-privileged information* conditions read:

“Both you and the other person evaluating this house will receive 4 pieces of information about the house. After reviewing this information, you’ll be asked to answer some questions about the house. The four features appear on the back of the cards below. We will turn over all four cards.”

Those in the *privileged information* conditions instead were told that this provision of information would be asymmetric:

“One of you—you or the other person evaluating this house—will receive 4 pieces of information about the house. The other of you will receive 3 pieces of information about the house. After reviewing this information, you’ll be asked to answer some questions about the house. The four features appear on the back of the cards below. For one of you, we will turn over all four cards. For the other of you, we will turn over a randomly selected three of the cards.”

These privileged information participants—like their non-privileged counterparts—received all four pieces of information. Based on the key feature manipulation, that key (privileged) feature—supposedly hidden from the other—was either the positive or the negative one.

Next, each card was turned over so that participants could learn the four features of the house. The order of such revelation was random, but subject to one constraint. Like in Study 3, the key feature—the one that was either positive or negative (and for privileged information participants, hidden from the other)—was always revealed last. That is, some saw the *positive* key feature last: “Solar panels on the roof keep the electricity bill low. Others saw the *negative* key feature last: “The kitchen has a shortage of cabinet and storage space.” This again holds constant the potential influence of a recency effect.

Baseline (factual) SEEs. To begin, all participants provided SEEs based on the information the yoked other was said to receive—either the complete set of four features (non-privileged condition) or the partial set of three features (privileged condition). Two items were used. One asked participants to predict how the yoked other would evaluate the house. This was made on a 101-point slider scale anchored at -50 (*very negatively*) and +50 (*very positively*). The

second item asked participants to estimate how likely the yoked other would be to recommend that users complete an on-site visit of the house. This percentage response was also offered on a 101-point slider scale anchored at 0% and 100%. After placing both items on the same (0-to-100) scale, these judgments were averaged to form a *factual evaluation* composite. Note that each factual evaluation was based on a different set of information (one that did or did not include the key feature). For this reason, these evaluation composites are not meant to be compared to each other. Instead, each condition's factual evaluation composite will serve as a comparison standard for a different (either additive or subtractive) evaluation composite. One comparison will test the PIE that each previous study has tested; the other, the subtractive variant on it.

Additive and subtractive SEEs. At this point—depending on their privileged information condition—all participants made additive or subtractive SEEs. Those in the privileged information condition completed a measure much like we have used before, one that asked participants to estimate how the other would evaluate the target home if the key feature *were* revealed to the yoked other. Note that this is an *additive* SEE: It asks how a target would be evaluated if a feature were added to the other's knowledge of the target. Of interest is whether these additive judgments are more extreme than the initial judgments offered by those who held non-privileged information. This tests the PIE: Does an attribute sway SEEs more when it has yet to be (vs. already has been) disclosed to another?

Those in the non-privileged information condition instead considered the complement of this possibility. That is, they considered how the other would evaluate the home if they *had not* learned the key feature. Note that this is a *subtractive* SEE: It asks how a target would be evaluated if a feature *were not* known by the other. Here, the key question is how these subtractive SEEs compare to the initial evaluations of those considering another who actually did

not learn about the key feature. This tests whether merely making information focal leads it to exert a disproportionate sway on SEEs, or whether this variant will reflect a boundary condition (as the segregated processing account anticipates).

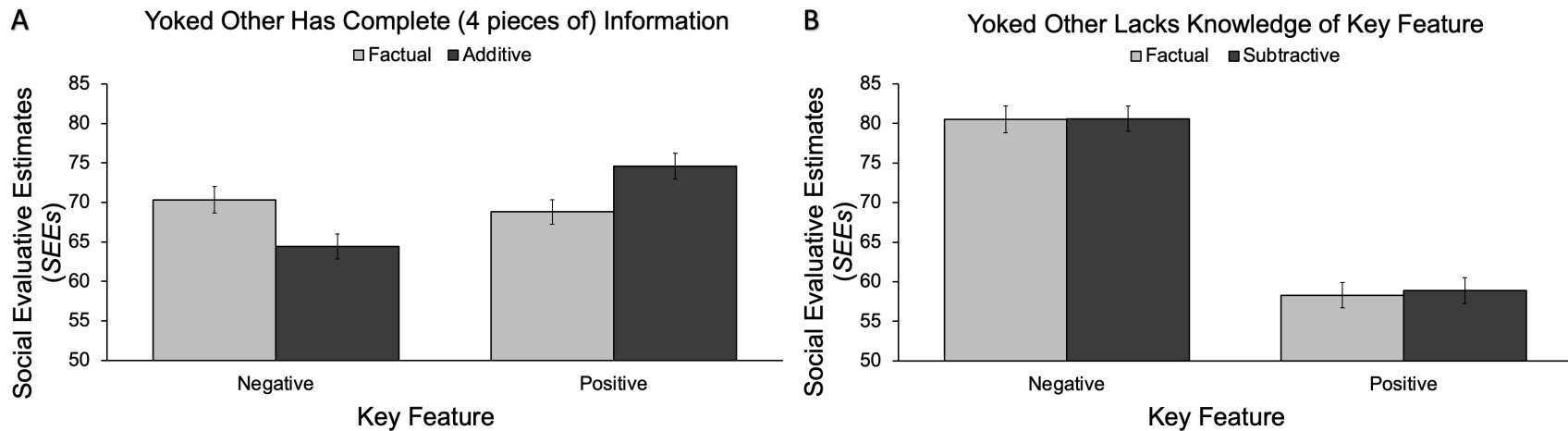
Results and Discussion

To begin, we considered those judgments about a target who either actually (non-privileged condition) or potentially (privileged condition) had complete information about the house. We submitted these forecasts to a two-way 2(information: privileged or non-privileged) X 2(key feature: positive or negative) ANOVA. Critically, we observed a significant Information X Key Feature interaction, $F(1, 461) = 13.00, p < .001, \eta_p^2 = .027$ (see Figure 5). Although the significant interaction reflects a replication of the PIE, we conducted pairwise comparisons to see whether the PIE emerged regardless of whether the key feature was positive or negative.

When the key feature was negative, privileged information participants thought that the yoked other would ultimately evaluate the job candidate more negatively ($M = 64.43, SD = 19.03$) than did non-privileged participants who did not think this feature was privileged ($M = 70.33, SD = 17.28$), $t(225) = 2.44, p = .016, d = 0.32$. Mirroring this result, when the self possessed positive information about the job candidate that the yoked other did not, these privileged information holders thought the other would ultimately evaluate the candidate more positively ($M = 74.61, SD = 15.12$) than did those non-privileged information participants who thought the feature was already common knowledge ($M = 68.80, SD = 18.09$), $t(236) = 2.67, p = .008, d = 0.35$. In summary, when information about the house—either positive or negative—was known by the self (but not a yoked other), participants believed that others' assessments of the house would ultimately be more positive or more negative, respectively, than when that information was not privileged.

Figure 5

Participants' SEEs about the Homebuying Opportunity, by Information Condition (Study 5)



Note. Panel A: Depicts SEEs of another who knew (factual) or was imagined to learn (additive) the key feature. The significant interaction reflects the PIE. Panel B: Depicts SEEs of another who actually lacked knowledge of the key feature (factual) or was counterfactually imagined to lack such knowledge (subtractive). The absence of an interaction is inconsistent with the pure focalism account. Error bars represent +/-1 SEM.

Does the PIE reflect the special influence of features that are imagined to be acquired by others, or does it merely reflect how making information unconditionally focal leads that information to have an outsized influence on SEEs? That is, does *subtracting* a feature that was already understood to be learned as one fact among many also lead SEEs to be particularly swayed? For this test, we examined SEEs about targets who knew only three of the four features of the house: either because the other target was said to have actually learned three of the four features (privileged information condition) or because participants had counterfactualized that this knowledge was never actually learned (non-privileged information condition). We submitted these SEEs to a two-way 2(information: privileged or non-privileged) X 2(key feature: positive or negative) ANOVA. As something of a manipulation check, we observed a large main effect of valence, $F(1, 461) = 152.71, p < .001, \eta_p^2 = .249$. That is, SEEs were more positive [negative] when the other supposedly lacked knowledge of the negative [positive] feature. But crucially, we did not observe an Information X Key Feature interaction, $F < 1$ (see Figure 5). This result is consistent with the segregated processing account: The PIE was specific to closing information asymmetries. That participants' SEEs were not similarly swayed by subtracting information that another already knew suggests that the PIE is not merely the result of information becoming unconditionally focal.

Study 6

Study 6 extended on the previous studies in two primary ways. First, we tested for the PIE with a new set of materials. Like participants in Studies 1 and 2, Study 6 participants developed SEEs about a real target. More specifically, participants learned about Move (shopmove.co), a San Francisco-based grocery delivery startup that promises users, “Award-winning staples—Made ethically—Sold at cost.” As one might anticipate, there is a catch to

gaining access to Move’s “radically fair prices”: a non-trivial yearly subscription fee. Move does not prominently display this to fee to potential subscribers. This—along with pretest validation (see Supplemental Materials) that this feature was in fact negative—served as a motivation for using this information as a (sometimes) privileged detail in Study 6.

Second, and of more central importance, Study 6 sought to more directly test whether segregated processing does indeed underlie the PIE. Studies 4 and 5 identified boundary conditions that were consistent with the segregated processing account (and not the pure focalism account), but these tests were indirect. That is, while we identified boundaries that were logically derived from the segregated processing account, such paradigms did not *directly* test the importance of how the privileged information was processed. Study 6 instead manipulated such processing directly.

In Study 6, some participants knew that they held privileged information, but we pushed some of these participants to process that information in a desegregated way. This experimental approach—steering people away from their natural information processing tendencies that are posited to explain judgment-and-decision-making effects—has been used to demonstrate the role of various information processing patterns in producing other biases in social judgment (Moon, Gan, & Critcher, 2020; Savitsky, Epley, & Gilovich, 2001). Although the effect we document and the mechanism we hypothesize underlies it are novel, we take an analogous approach. More specifically, we asked participants in a new *privileged-desegregated* condition to consider that the yoked other would eventually learn all of the information, but we had them reprocess this privileged information in an integrated manner. More specifically, we had them: 1) rank the extent to which each feature (including, crucially, the privileged one) would contribute to a positive evaluation of the target and 2) write about how each individual feature would contribute

to the other's attitude toward the target. Note that these tasks require privileged-desegregated participants to construct an SEE not simply by adding the privileged attribute, but by reconstructing the SEE to allow the privileged attribute to be one among many pieces of information that define the target. In the interest of maximizing internal validity, we also had participants in both our non-privileged and (standard) privileged conditions complete ranking and writing tasks. But crucially, privileged information participants did so by considering only the information that the other already knew (thereby leaving the privileged information segregated).

Study 6 thus has two central predictions. First, we expected to replicate the PIE by showing that those holding privileged (negative) information would estimate that the other would ultimately evaluate Move more negatively than those for whom this information was never privileged. Second, we expected that by encouraging privileged information holders to process the privileged information in a desegregated manner, that the size of the PIE would be reduced. That is, we expected that the SEEs of those in the privileged information condition would be more negative than those in both the non-privileged and privileged-desegregated conditions.

Method

Participants and design. One thousand thirty-three Americans were recruited from AMT in exchange for nominal compensation. Participants were randomly assigned to one of three information conditions: *privileged*, *privileged-desegregated*, or *non-privileged*. Forty-two participants failed a memory-based attention check (see Supplemental Materials) and were excluded from all analyses reported below. This left a final sample of 991 participants. Our hypotheses, methods, exclusion criteria, and analytic plan were preregistered:

https://osf.io/t7dfp/?view_only=d7ba950c8cd6495bb929d1948e6eb90e.

Procedure. Participants initially learned that they would be evaluating a grocery delivery service called Move. Participants read how Move—an “online supermarket” startup that launched in 2019—offers members the ability to purchase ethically sourced grocery products at reasonable prices. Participants learned that Move was conducting marketing research to better understand consumer preferences. To aid with this endeavor, participants learned that they would be evaluating several of Move’s features:

“Move wants two separate people to evaluate different features of their grocery delivery service. This means that you will be matched with an Amazon Mechanical Turk participant also taking the survey right now.”

Next, participants read that both they themselves and the yoked other would receive information about several of Move’s features. As in previous studies, all participants saw four cards; each card detailed one of Move’s features. When the cards were initially presented, the four features themselves were hidden from participants. Participants in the *non-privileged information* condition read that:

“Both you and the other participant evaluating Move will receive 4 pieces of information about the grocery delivery service. After reviewing this information, you’ll be asked to answer some questions about Move. The four features appear on the back of the cards below. We will turn over all four cards.”

In the two *privileged information* conditions, participants learned that this information would be distributed asymmetrically:

“One of you—you or the other participant evaluating Move—will receive 4 pieces of information about the grocery delivery service. The other of you will receive 3 pieces of information about the grocery delivery service...”

Then, to amplify the realism of the matching paradigm, all participants were directed to a screen with a loading GIF. After waiting for a few seconds, participants received a message communicating that they had (ostensibly) been matched with another participant currently completing the study.

Next, each card containing one of Move's features was turned over. All participants—those assigned to either the privileged or non-privileged conditions—received all four pieces of information about Move. In a pretest ($N = 102$ American AMT workers, see Supplemental Materials for complete details), participants evaluated ten features. Three of those features were evaluated (similarly) positively and became the three positive feature (“For every order, free 2-day nationwide shipping is guaranteed.”; “For customer service, consumers can chat with a real person.”; “Shoppers can see the region of the world where each product is produced.”), while one feature that was evaluated negatively became the negative feature (“The membership fee is \$195/year.”). Participants in the non-privileged information condition learned that the other participant was also seeing all four of Move's features. Conversely, participants in the two privileged information conditions learned that the other participant was not seeing the negative feature. As in Studies 3 and 5, for all participants, three cards were turned over in a random order, and then this negative feature was always revealed last.

Ranking Task. Next, participants saw a graphic that summarized the information about Move that they themselves and the other participant had learned. This meant that all participants were shown the four features in parallel. In an adjacent column, participants saw those features that the other participant had learned. For the non-privileged information condition, this column merely repeated all four features. For those in the privileged information conditions, this column was lacking the one privileged feature.

Accompanying this graphic were instructions that asked participants to rank-order how the other participant would evaluate Move's features. That is, participants were asked to drag and drop Move's features according to how positively the other participant would evaluate them (i.e., "the feature that you think the other participant will evaluate most favorably will be #1"). Critically, the number of features participants rank-ordered depended on the information condition to which they were assigned. In the *non-privileged* information condition—as the other participant had ostensibly seen all four of Move's features—participants rank-ordered all four features. In the *privileged* information condition—in which the other participant had supposedly seen only three of Move's features—participants rank-ordered only these three features. But in the *privileged-desegregated* condition—in which the fourth (negative) feature had not yet been shown to the other—participants considered how the yoked other would rank-order all *four* features upon learning all of them. This was the first of two steps designed to encourage some of those holding privileged information to achieve a desegregated perspective on how the other would (eventually) construe the target.

Writing Task. The writing task was designed to reinforce the ranking task. More specifically, participants were asked to write—in an open-ended text box—a brief statement about how the yoked other would evaluate Move in light of its features. They were told they should reference each of the features individually (i.e., "In light of Feature A, they would probably...") Those in the non-privileged and privileged information conditions again considered how the other would evaluate Move in light of either all four features (non-privileged) or the three known features (privileged). Those in the privileged-desegregated condition were asked to consider how the other would evaluate Move upon learning all four pieces of information. Thus, once again, those in the privileged-desegregated condition were: 1)

reminded that the other did not currently hold this information, but 2) encouraged to develop a perception of how the privileged attribute would operate as one among many pieces of information that would guide the yoked other's attitude toward the target. Participants were required to spend at least 60 seconds on this task.

SEEs. On the next page, participants completed the main dependent measures that asked how the yoked other would evaluate Move. As in previous studies, the exact form of these SEEs varied by condition. Participants in the non-privileged condition—who assumed that the yoked other already had complete information about Move all along—estimated on a slider scale how the yoked other would evaluate Move (from *very negatively* [-50], to *very positively* [+50]) and how interested the yoked other participant would be in purchasing a Move membership (from *not at all interested* [-50], to *extremely interested* [+50]). Participants in the privileged information condition were asked to “imagine that the other participant did learn about Move’s Feature D.” As in previous studies, we added the words “If so:” before the dependent measures. Participants in the privileged-desegregated condition completed the same measures, but ones which instead led with “Still imagining that the other participant did learn about Move’s Feature D.” In this way, participants in all three conditions were making SEEs about another who held complete information about Move. The conditions varied in whether that information was known all along (non-privileged) or potentially knowable (privileged, privileged-desegregated). But crucially, privileged and privileged-desegregated participants varied in whether participants had been encouraged to process the final feature as one among all four features (privileged-desegregated) or as a feature that—until then—had been segregated from the rest (privileged). The two SEE measures were averaged to form an *evaluation* composite ($r = .66$).

Results

To begin, we tested whether participants' SEEs differed between the three information conditions. And indeed, this was the case, $F(2, 988) = 24.56, p < .001, \eta_p^2 = .047$ (see Table 2). We proceeded to conduct a series of pairwise comparisons that could address our central hypotheses. First, we tested whether we would replicate the privileged information effect (non-privileged vs. privileged). Second, we examined whether encouraging participants to incorporate the privileged feature in a more integrated, desegregated form reduced its influence on SEEs (privileged vs. privileged-desegregated). Third, we tested whether this intervention eliminated some or all of PIE (non-privileged vs. privileged-desegregated).

First, we replicated the PIE: Participants in the privileged information condition believed that the yoked other would evaluate Move more negatively ($M = 2.41, SD = 25.63$) than did participants in the non-privileged information condition ($M = 15.02, SD = 20.25; p < .001$), $t(988) = 7.01, p < .001, d = 0.55$. Second, we found that participants encouraged to desegregate the initially privileged piece of information estimated that the yoked other would be more positive about Move ($M = 8.59, SD = 22.90$) than those in the standard privileged information condition, $t(988) = 3.44, p < .001, d = 0.25$. And third, we saw that the desegregation intervention eliminated some, but not all of the PIE, $t(988) = 3.63, p < .001, d = 0.30$.

Discussion

In summary, Study 6 provides evidence that the PIE is—at least in part—a function of the segregated information processing that holding privileged information encourages. Although participants in the privileged-desegregated condition knew that one of the pieces of information was initially withheld from the other participants, they were encouraged to more fully integrate that information into their understanding of how the other would (eventually) view the target. More specifically, through a combination of ranking and verbal elaboration, such participants

Table 2*SEEs of Yoked Others' Evaluations of Move, by Information Condition (Study 6)*

SEE	Information Condition		
	Non-privileged	Privileged	Privileged-desegregated
Evaluation	23.12 (18.36) _a	6.32 (26.97) _c	15.19 (23.75) _b
Purchase Interest	6.92 (25.85) _a	-1.50 (28.80) _b	2.00 (26.99) _{ab}
<i>Evaluation Composite</i>	15.02 (20.25) _a	2.41 (25.63) _c	8.59 (22.90) _b

Note. Means (with accompanying standard deviations) in the same row that do not share the same subscript differ at the $p < .01$ level.

were prompted to process that information in a desegregated fashion, as one piece of information among many. Although these findings support the crucial role of information segregation in the PIE, the results are ambiguous as to whether this feature explains only some or all of the PIE.

That is, the desegregation manipulation may have eliminated roughly half of the effect because the PIE is multiply determined. Alternatively, the manipulation may not have *completely* desegregated the information. After all, even privileged-desegregated participants may have begun to form an understanding of how the other would view the app just in light of the three pieces of information before our interventions began. Even as these results indict the crucial role of segregated processing in producing the PIE, they do not test *how* privileged information exerts a heightened effect on SEEs. Study 7 examines this more directly.

Study 7

Study 7 extends on our previous studies in three ways. First, we moved to a new privileged information context. Participants were asked to consider two single people they knew whom they might want to set up. We provided participants with an email template which would guide them through the process of introducing one single person (“the dating prospect”) to the recipient of the email. The template required participants to provide four pieces of information to the email recipient about the dating prospect: three selling points (positive features) and one potential reservation (negative feature). Whereas *non-privileged* information participants were told that the entire email would be sent to the recipient, *privileged* information participants were instead shown a template of the email in which the one reservation had been edited out. The former participants estimated how much romantic interest their full email would inspire. The latter participants estimated how much romantic interest would be inspired if the full email were sent instead. This paradigm builds on our previous efforts in several ways. Whereas Study 2

participants themselves did elaborate on experimenter-provided information (about a technology product they had just tried) that would then be supplied to another participant, Study 7 went a step further by having participants themselves generate all of the information about the target that would be shared with another. Furthermore, whereas participants in our previous studies made SEEs about an unknown yoked other, participants in Study 7 were personally acquainted with the person whose evaluations they would be estimating.

Second, participants were not asked to consider how another would respond if they eventually learned the fourth piece of information. Instead, participants considered the past possibility that the email they sent *did* contain this fourth piece of information. In this way, participants were clearly not offering SEEs about a target who would learn this information in two stages; rather, they made SEEs about a target by considering the possibility that they would learn all four details all at once (instead of only just three). On the one hand, the typical question that real-world privileged information holders ask themselves is how someone else would respond upon learning such information, which means our previous paradigms that asked such a question may have been higher in external validity. Study 7 makes this change because (beyond the evidence in Study 3 that the PIE is an error) it would allow us to more definitively conclude that any difference in SEEs between conditions—i.e., the PIE—reflects a normative error. In the General Discussion, we more broadly consider the (sometimes subtle) variety of measurement approaches we have used to detect the PIE.

Third, we wanted to more directly test for *how* privileged information disproportionately colors SEEs. That is, we have argued that when privileged information is incorporated in a segregated fashion, it is not incorporated on equal footing with the other information about the target. As a result, it should be *weighted* more heavily in the formation of an SEE (compared to

when the information is not privileged). A cognitive algebra approach to person perception has long appreciated that impressions of other people are not a simple sum of the impressions of their individual parts, but instead reflect a weighted average whereby some dimensions serve as stronger and thus more important input than do others (Anderson, 1968; Himmelfarb, 1973). We propose that impressions of others' attitudes should be characterized by the same property, with predictably enhanced weighting of privileged information.

But by an alternative attribute extremity possibility, it may not be that privileged information exerts a stronger influence on SEEs, but instead that the very meaning of or perceived extremity (i.e., the assumed positivity or negativity) of information shifts when it is privileged. This idea is made plausible by previous research that has found that keeping information hidden can lead people to extremify that information's valence (Lane & Wegner, 1995; Wegner, Lane, & Dmitri, 1994). Secrets often beget rumination (Slepian, 2021), and rumination can lead information to seem more extreme (Tesser & Paulhus, 1978; Lyubomirsky, Tucker, Caldwell, & Berg, 1999). That is, when the potential reservation is privileged (as opposed to common knowledge), the email writer might inflate the extent to which the recipient would see it as a "red flag."

More generally, these two mechanisms would fall within the "Weighting" and "Assessment" components of Rosenzweig and Critcher's (2014) SAW model. The enhanced *weighting* account suggests that a shift in regression betas will show that SEEs are more a function of information when it is privileged than when it is common knowledge. The attribute extremity account suggests that the valence of privileged information may be differentially *assessed* by privileged and non-privileged information holders. Our final study tests whether one or both of these accounts explain the PIE.

Method

Participants and design. Two hundred thirty-one Americans were recruited from AMT in exchange for nominal compensation. Participants were randomly assigned to one of two information conditions: *privileged* or *non-privileged*. Forty-one participants failed a memory-based attention check (see Supplemental Materials) and were excluded from all analyses reported below. This left a final sample of 190 participants.

Materials and procedure. Participants learned that they would be taking part in a project called *Make a Friend, Or Something More!*, in which they would have the opportunity to set up two people whom they knew. All participants read the following information:

“As part of this project, you will be asked to think about someone that you know well and care about (e.g., a friend or family member) who is single. We will call this person your ‘close other.’ You will think about a person—whom you also know well—whom you might set your friend or family member up with. We will call this person the ‘dating prospect.’ *Make a Friend, Or Something More!* will help you compose an email letting your friend or family member know about the person you might want to set them up with. For example, *Make a Friend, Or Something More!* could help you suggest to your sibling (your close other) that s/he go on a first date with your coworker (the dating prospect).”

Next, participants were prompted—in two separate text boxes—to enter the first names of the close other and the dating prospect. After advancing the screen, participants then read about how *Make a Friend, Or Something More!* would walk them through the process of writing an email that would introduce the close other to the dating prospect. Further, we told participants that—with their permission—this email would actually be sent to the close other. To help to explain the task, participants next saw an (incomplete) email template that they would complete as part

of *Make a Friend, Or Something More!* The template prompted participants to fill in four pieces of information about the dating prospect:

“Hi [close other],

Hope that you are doing well! I’m reaching out because I am participating in a project that encourages us to let a friend or family member who is single know if we know of someone they might be interested in going out with. I know that it’s a bit weird to send an email out of the blue, but am giving this a try. I hope that you don’t mind, and obviously it’s completely up to you what you want to do.

Ok, I wanted to tell you a little bit about [the dating prospect]. I’ve known [the dating prospect] for a little while, and I think that the two of you would get along great...

Just to tell you a little bit about [the dating prospect]:

- {A first fact about [the dating prospect]}
- {A second fact about [the dating prospect]}
- {A third fact about [the dating prospect]}
- {A fourth fact about [the dating prospect]}

Ok, let me know if you are interested and I can set the two of you up by text or email!

Hope that things are going well and talk soon.”

Next, participants were tasked with detailing four facts about the dating prospect that could be shared with the close other. Participants read a series of instructions that detailed how the first three facts that they shared would be facts about the dating prospect that should trigger the close other’s interest. After all, participants were writing this email to someone they thought would be a good match for the dating prospect, so we expected that this would be relatively simple. In contrast, the fourth and final fact was supposed to be something a bit more negative,

the sort of thing that might complete the sentence, “Of course, no one is perfect. Really, the only thing I can think of that’s not ideal is...”. In this way, all participants supplied three positive or flattering features as well as one negative or less flattering feature. Whether all of these facts would actually make it into the to-be-shared email varied by *information* condition.

Participants in the *non-privileged* information condition learned that all four of the facts about the dating prospect would be included in the email to their close other. Those in the *privileged* information condition learned instead that only three of those facts—those that highlighted the dating prospect’s flattering, favorable qualities—would be shared in the email. This established an information asymmetry for privileged information participants. Participants in this privileged information condition were eventually asked to consider how their close other would evaluate the dating prospect if the close other instead received the full email that did not redact the (unflattering) feature.

Participants then completed two items that assessed how they believed the close other would evaluate the dating prospect. The first item asked how the dating prospect would be evaluated overall. This item was on a 101-point slider scale anchored at -50 (*very negatively*) and +50 (*very positively*). The second item asked participants to estimate how likely it would be that the close other would reply to the email to express interest in being introduced to the dating prospect. This response was given on a 101-point slider scale anchored at 0% and 100%. After transforming the first item onto a 0-to-100 scale, the two items were averaged to form a *general evaluation* composite ($r = .71$).

Participants in the non-privileged information condition completed these items only once, but with the knowledge that the close other would receive the email with all four pieces of information about the dating prospect. Participants in the privileged information condition

completed these items twice: once based on how they expected the close other would respond in light of the email (which omitted the negative feature), and then again under the alternative possibility that the email had included all four pieces information. In the privileged condition, only this second set of judgments was used in all analyses.

Next, participants considered the four features one at a time. For each, participants estimated how the close other would respond to that particular fact. Each judgment was made on a 101-point slider scale. These were anchored at -50 (*very negatively*) and +50 (*very positively*).

Results and Discussion

We began by testing for evidence of the privileged information effect. As hypothesized, close others were estimated to evaluate the dating prospect more negatively in the privileged condition ($M = 68.19, SD = 20.23$) compared to the non-privileged condition ($M = 74.27, SD = 15.14$), $t(187.43) = 2.37, p = .019, d = 0.34$. In other words, the email writers estimated that their readers would have a less negative evaluation of the dating prospect if the negative fact were actually included in the email, compared to if the email writers were merely considering the counterfactual possibility that the full email that included the negative detail were sent.

Next, we proceeded to test which of the considered mechanisms explained the PIE. Toward this end, we defined two variables that drew on participants' estimates of how the four facts—when considered individually—would be evaluated by their close other. One was the *target feature evaluation*. This reflected participants' estimates of how the close other would evaluate the less flattering feature in isolation—the one that was privileged for some participants. The second was the *alternative features evaluation*. This reflected an average of the other three estimates—those that corresponded to the dating prospect's flattering features—which were always communicated in the email itself.

How did participants estimate the close other would evaluate the individual features? We submitted participants' feature evaluations to a 2(information condition: privileged or non-privileged) X 2(feature evaluation: target or alternative) mixed-model ANOVA, with only the first factor manipulated between subjects. We found an unsurprising main effect of feature evaluation, $F(1, 188) = 64.70, p < .001, \eta_p^2 = .256$, indicating that participants believed that close others would evaluate the unflattering feature of the dating prospect less positively than the alternative, flattering features. We observed neither a main effect of information condition, $F < 1$, nor a significant interaction between information condition and feature evaluation, $F(1, 188) = 2.42, p = .121, \eta_p^2 = .013$. In other words, possessing privileged information did not change how participants thought the targets' individual features would be evaluated. This is inconsistent with the attribute extremity account.

Were privileged and non-privileged participants' SEEs equally sensitive to the target and alternative features? Or instead—consistent with the enhanced weighting account—were privileged SEEs (compared to non-privileged SEEs) more closely tied to evaluations of the target feature? We tested this by regressing the general evaluation composite on participants' information condition (-1 = non-privileged, +1 = privileged), the target feature evaluation (standardized), the alternative features evaluation (standardized), as well as the interaction between each feature evaluation composite and the information condition.

Replicating the existence of the PIE, this revealed a main effect of information condition, $\beta = -0.14, t(184) = 2.77, p = .006$. Further, we also observed main effects of the target feature evaluation, $\beta = 0.36, t(184) = 7.06, p < .001$, and the alternative features evaluation, $\beta = 0.54, t(184) = 10.55, p < .001$. Both indicate that participants' general evaluations were sensitive to how they expected their close other to evaluate the dating prospect's individual features.

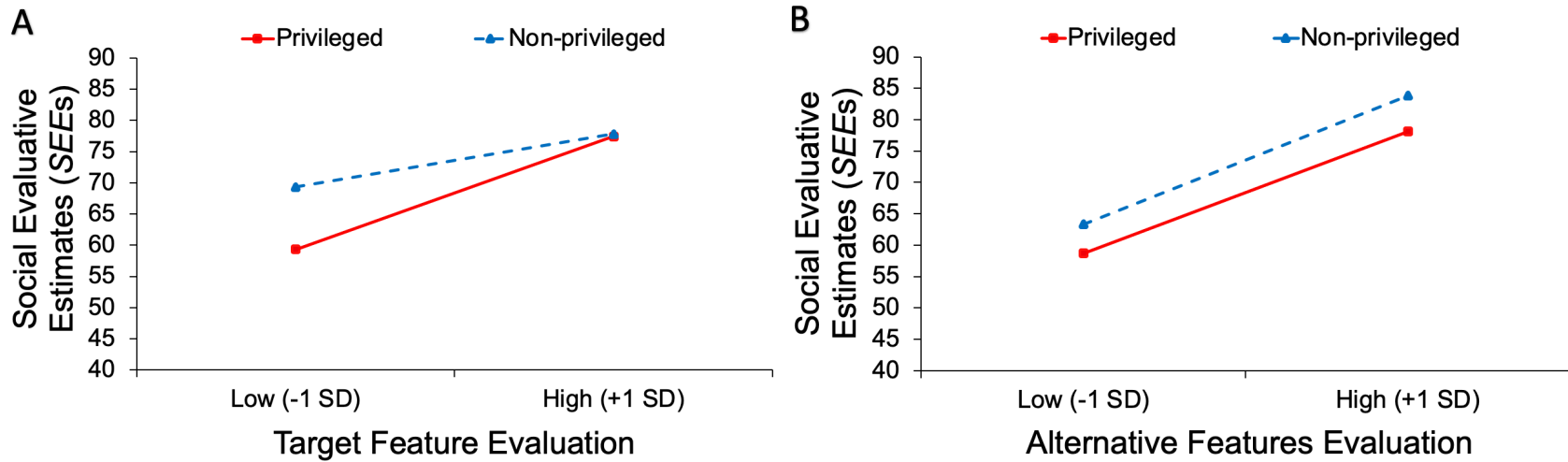
We next turn to the interaction terms. There was no Information Condition X Alternative Features Evaluation interaction, $\beta = -0.01$, $t < 1$. We did, however, observe a significant Information Condition X Target Feature Evaluation interaction, $\beta = 0.13$, $t(184) = 2.55$, $p = .012$ (see Figure 6). When the unflattering fact about the dating prospect was privileged, evaluations of this feature strongly informed participants' general evaluations, $\beta = 0.49$, $t(184) = 6.92$, $p < .001$; when the unflattering feature was common knowledge, this effect was reduced, $\beta = 0.23$, $t(184) = 3.14$, $p = .002$. This suggests that the PIE can be localized to privileged information being weighted more heavily in SEEs.

General Discussion

Economists, political scientists, sociologists, and social psychologists have long studied how information asymmetries can affect consumer markets (Akerlof, 1970; Healy & Palepu, 2001; Stigler, 1961; Lewis, 2001), enable moral hazards (Schulz, Lubatkin, Dino, & Buchholtz, 2001; Stiglitz, 2002), limit effective communication in both close relationships (Blau, 1964; Stanley, Rhoades, Scott, Klemer, Markman, & Fincham, 2016) and negotiations (Thaler, 1988; Becker, Clement, & Noth, 2016; Maaravi & Levy, 2017), and even stunt the efficacy of policy (Welker, 1995; Tirole, 2015). Much of this existing work has highlighted how people operate egocentrically when considering *open* information asymmetries (e.g., Camerer et al., 1989). The present work distinguishes itself by examining a complementary question: Do people who possess privileged information assume that information will sway another's evaluations more upon the potential *closure* of the information asymmetry (compared to when no such information asymmetry has existed)? That this question has received little attention in existing literature is surprising. After all, in many cases secrets are ultimately revealed, classified information becomes declassified, and privileged knowledge becomes common knowledge (Dunbar, 2004;

Figure 6

General Evaluation Composite and Target and Alternative Feature Evaluations (Study 7)



Note. Panel A: Model-predicted general evaluation composite by the information manipulation and the estimated evaluations of the target feature (at 1 SD above and below the mean). Panel B: Model-predicted general evaluation composite by the information manipulation and the estimated evaluations of the alternative features (at 1 SD above and below the mean).

Dubar, Marriott, & Duncan, 1997; Slepian, Chun, & Mason, 2017; Slepian & Greenaway, 2018; Slepian & Kirby, 2018).

Across seven studies, we documented and shed light on what gives rise to a *privileged information effect* (PIE): People assume that information will affect others' evaluations more when people add (initially privileged) information to their understanding of what another knows compared to when that information was simply always part of what another was understood to know. In what follows, we review the totality of our evidence for the PIE before situating our findings within broader literatures on asymmetric information, disclosure decisions, and secret keeping.

Summary of Empirical Results

The robustness of the privileged information effect. First, we identified the PIE and established its robustness. Some paradigms used more naturalistic materials: variations of an actual online advertisement (Study 1), participant-supplied reviews of a new technology product they tried (Study 2), and a dating “set-up” email that participants composed to a known other (Study 7). Others used a more stripped-down card-turning paradigm that sacrificed some external validity to more fully standardize participants' experience in receiving (sometimes privileged) information about an investment opportunity (Study 3), a job candidate (Study 4), an on-the-market home (Study 5), or a real-life startup (Study 6). The PIE emerged consistently across the contexts, regardless of whether the privileged information was negative (Studies 1-2, 4-7) or positive (Studies 3, 5). The PIE did not emerge due to an implicit signal that the privileged information was in need of concealment: It emerged even when the privileged feature was said to be chosen at random (Studies 3-5) and even when the experimenter had no preexisting knowledge of the privileged attribute (Study 7). In the latter case, the privileged information

manipulation could not have been a commentary on the attribute's extremity (a possibility that Study 7's results also directly ruled out). Further, we observed the PIE when information was generated and shared by the participants themselves (Studies 2, 7), supplied exclusively by a third party (i.e., the experimenter; Studies 1, 3-6), and even when the participants themselves knew the person with whom privileged information was not being shared (Study 7).

The privileged information effect as a bias. Second, we asked whether the PIE is a *bias*, and relatedly, whether it is a bias that has some truth to it. Participants who learned (negative) information that was initially withheld from them *did* evaluate a target more negatively than did those who knew the negative information all along (Study 3). That said, privileged information holders clearly overestimated how much this (initially) privileged information would ultimately sway others' evaluations, and thus were biased. However, Supplemental Study A was unable to replicate the finding that initially withheld information is more influential on recipients' evaluations of a target. More generally, support for the PIE appears to be clear and robust, but recipients of privileged information may only *sometimes* be disproportionately moved by it. Furthermore, Study 7 offered the cleanest standard for establishing the PIE as a bias. Privileged information participants imagined that the yoked other had actually learned the (actually privileged) information along with the common knowledge. That the PIE emerged in this paradigm underscores that it is a normative error in social judgment.

Support for the segregated processing account over alternatives. Finally, we provided support for our argument that the PIE emerges because privileged information is processed in a segregated manner and is thus incorporated as a standalone piece of information that is weighted more heavily in overall SEEs. In five ways, we identified boundary conditions and mechanistic

evidence consistent with the theoretical logic of the segregated processing account, while also ruling out alternative accounts. First, the PIE is a distinctly social phenomenon: Privileged information holders' SEEs—but not their own evaluations—were disproportionately swayed by privileged information (Study 4). That is, it was not simply the case that having participants focus on a single piece of information leads it to be more top-of-mind and thus more heavily influential in *all* judgments. Second, Study 5 compared two circumstances in which participants were asked to psychologically manipulate a single piece of information when considering SEEs. When participants *added* new information that was first privileged to a representation of how another evaluates a target, the PIE emerged. But when participants *subtracted* a piece of information which was common knowledge, that information was not believed to have been similarly impactful. When information is added to another's representation, it is initially processed in a segregated fashion (and is thus disproportionately influential); when information is subtracted from another's representation, it is first processed as merely one among many attributes that characterizes a target (and thus its removal is not assumed to be similarly impactful). Taken together, Studies 4 and 5: (1) demonstrate that the PIE is not merely a consequence of information being unconditionally focal, and (2) identify boundary conditions predicted by the segregated processing account.

Third, it is worth emphasizing that the PIE does not emerge because of an anticipated recency effect. That is, when an attribute is privileged, then it would be—if one ponders the potential closing of an information asymmetry—imagined to be learned last. More recently learned information can disproportionately color evaluations of targets (Jones & Goethals, 1972; Parducci, 1964). However, the paradigms in Studies 3 and 5-7 either provided or had participants generate the piece of information that was (only sometimes) privileged last. In this way,

regardless of whether an information asymmetry was present, social evaluative estimates were made about a target who had learned (or was imagined to learn) this attribute last. Fourth, Study 6 had privileged information holders reconsider the privileged piece of information in light of the full set of information about the target. When the information was (re)processed in this desegregated manner, the size of the PIE was substantially reduced. Fifth, Study 7 distinguished between two means by which privileged information exerts an undue influence. In contrast to what findings in the secrecy literature might anticipate (for example, Lane & Wegner, 1995), we found no evidence for an attribute extremity account of the PIE: Privileged information holders did not assume that the privileged information would be interpreted in a more evaluatively extreme manner by recipients. Instead, privileged information was weighted more heavily when it was incorporated as (segregated) privileged information instead of information that had been common knowledge all along.

Generality of the PIE. We observed consistent evidence in support of the PIE both in samples of American undergraduates as well as among online samples drawn more broadly from the United States. That said, questions remain regarding whether the PIE would emerge in other cultural contexts as well. On the one hand, this might seem to be a question of whether there are predictable cultural differences in the extent to which people engage in more segregated and integrated processing. And indeed, there is some evidence that those from more interdependent cultural contexts show evidence of more holistic (and thus less segregated) processing, even for quite low-level perception and categorization tasks (McKone et al., 2010; Varnum, Grossman, Kitayama, & Nisbett, 2010). That said, the key question here is whether a predisposition to engaged in segregated processing is something that individual research participants brought with them to our studies (in which case we would expect cultural differences in such processing styles

to be potentially influential), or whether segregated processing is simply encouraged by the dictates of the judgment task itself. That is, privileged information holders naturally differentiate their understanding of what someone else currently knows from what they could know. In such cases, individual differences in the propensity to engage in more integrated processing may have little effect. Of course, these questions could be addressed most directly by testing for the PIE's generality across cultural contexts as well as across individuals whose cognitive styles are known.

Temporal Dynamics of the PIE

The resolution of information asymmetries. As privileged information holders consider the closure (or absence) of an information asymmetry, one can—as our studies did—guide participants to do that in different ways. In some of our studies, we asked participants to consider the future possibility of an information asymmetry closing (e.g., “if they were to learn...”; Studies 1-2, 4). In Study 7, participants considered the alternate reality that the yoked other never experienced an information asymmetry. The other three studies (3, 5-6) used language that could be interpreted either way; participants were asked to consider if the yoked other “did learn” the (previously privileged) information. These different operationalizations have different strengths. Considering the future closure of an information asymmetry mirrors how we suspect privileged information holders form SEEs about lesser-informed others' evaluations (“If they were to learn about this, I wonder...”). Conversely, considering instead the past reality that the information asymmetry had never existed offers a more conservative test; under this operationalization, any emergence of the PIE is necessarily a normative error. Notably, both operationalizations (or interpretations) require privileged information to be incorporated into SEEs in a segregated manner. That said, we conducted Supplemental Study B (N = 804 Americans from AMT) to test,

in a single experiment, whether the PIE emerged similarly in both cases. Although we invite interested readers to consult the Supplemental Materials for complete information, we both replicated the PIE and found that it emerged similarly in size regardless of the two operationalizations. This reinforces both the robustness of the PIE and our argument that the segregated processing of privileged information—as opposed to one singular description of how an information asymmetry closes—underlies the PIE. Furthermore, the similarity in size of the two operationalizations of the PIE suggests that (at least in our paradigms) what is driving the PIE is not a forecasted shock or surprise about later learning information one first did not know.

The longevity of the PIE. Our explanation for the PIE focused on how privileged information was processed and incorporated into an overall SEE in a segregated fashion. In fact, we argued that the state of a privileged information holder—who is simultaneously interested in what another *does* know as well as what another *could* know—only helps to reinforce this segregation. But what happens once these dual SEEs are no longer relevant—i.e., once the information asymmetry that was once open is long closed? Study 6—in which privileged information holders were prompted to process (even currently) privileged and non-privileged information in an integrated manner—certainly shows that the strength of the PIE is not unconditional. But in the real world, there are not experimental instructions that nudge such integrated processing. In other words, Study 6’s findings more support the mechanistic importance of non-integrated (segregated) processing in giving rise to the PIE than they do reveal when such integrated processing occurs naturalistically.

The longevity of the PIE thus presumably depends on whether an SEE is formed (or formed anew) in the moment, or whether it is directly recruited from memory. Once an information asymmetry closes, a (former) privileged information holder may simply recall the

SEE that he developed while holding privileged information (“I remember thinking this product was going to be a flop once people started using it”), an SEE that exaggerated the likely role of the privileged information. Such direct recall would suggest the PIE would persevere even after an information asymmetry has closed. In other cases, perhaps because the originally formed SEE is not remembered or because the (former) privileged information holder thinks the SEE merits a more careful inspection, the SEE may be formed anew. In such cases, there may not be reason to segregate the once-privileged information from the never-privileged information, and each piece of information—once-privileged and never-privileged alike—might occupy a similar, “one piece of information among many” status in forming this new SEE. We suspect that hiring committee members who know that they alone know about a job applicant’s felony are in a quite different position from those who know that there was an 18-hour period a few weeks ago when only they themselves had this (privileged) information. It is unclear why this felony conviction’s once-privileged status (for part of a day some weeks ago) would presently encourage the segregated incorporation of such information as they look around the room and consider what their fellow committee members think. That is, keeping in mind the distinction between what another *presently* knows versus could know clearly has more relevance, say, than holding dual representations of what another knew 13 days ago as opposed to 12.25 days ago.

Intentional Information Asymmetries

Secrecy and focalism. Although we repeatedly found evidence consistent with a segregated processing instead of a pure focalism account, we suspect that there are cases in which the pure focalism mechanism would also come into play. That is, certain types of privileged information may be particularly accessible—and thus, disproportionately influential—in people’s minds (Higgins, 1996). Holding secrets—and thus intentionally or strategically

prolonging an information asymmetry—may be one such case. Existing work has suggested that holding secrets can make them loom large in the self’s mind (Maas, Wismeijer, Van Assen, & Aquairius, 2012; Quinn & Chaudoir, 2009; Slepian, Kirby, & Kaloberinos, 2020; but see also Slepian, Greenaway, & Masicampo, 2020). Further, intensely personal secrets that are potentially shaming—such as one’s hidden stigmatized identity—may be particularly top-of-mind and more generally influential (see, for example, Quinn & Chaudoir, 2012).

Yet, one empirical challenge lies in differentiating whether it is the secrecy itself that *causes* this information’s chronic accessibility, or whether some third variable (e.g., the perceived importance or valence of the secret) that causes it to be both chronically accessible and kept hidden from others. Naturalistically, this might produce a pattern that mirrors the PIE: Those who imagine what it would be like to close an information asymmetry (and reveal their closely guarded secrets) might imagine more extreme evaluative reactions than those who have already closed such asymmetries (because they did not feel the need to keep the information secretive to begin with). Of course, such an effect would speak more to confounded differences in the type of person who *maintains* information asymmetries as opposed to the effects of holding privileged information per se. A person who keeps their cannabis consumption a secret likely thinks others will judge them more harshly than does a person who more openly discloses this same information in conversation. Our own studies avoided this confound by randomly assigning participants to experience (or not) information asymmetries. Although this approach allowed us to establish the causal role of holding privileged information in producing the PIE, we certainly do not exclude the possibility that there exist contexts in which additional mechanisms (e.g., focalism) may explain the PIE or related phenomena.

Having information withheld from the self. There are many cases in which better-informed parties might have the desire to strategically maintain information asymmetries. Think back to the cases of Oculus or Cabin. In navigating the terrain between recklessly advertising and blatantly withholding information about their products' flaws, both firms are *prolonging* an open information asymmetry. At first glance, many consumers are likely unaware that Cabin is a bus service or that the Oculus Go allows light to seep through the bottom of the headset. But of course, these information asymmetries will eventually close. And when they do, privileged information holders—and particularly those who are seen as prolonging information asymmetries—may be judged negatively.

Most research in this vein has focused on cases in which there is some *known unknown*: That is, an individual is (known) to be withholding some (unknown) information. When information about a target is knowingly withheld, judges often assume that information is particularly damaging (John, Barasz, & Norton, 2016). On top of this, privileged information holders who dodge questioning about such known unknowns may be judged particularly negatively (Rodgers & Norton, 2011). But consider that information asymmetries are often closed via the reveal of some (previously) *unknown unknown*: information—such as the Oculus's loose nose bridge—that lesser-informed others' do not even *know* that another party is withholding.

In such cases, those who lack information will not know information is being withheld until such information is revealed. Thus, prior to such a reveal, individuals kept in the dark do not even know the dimension on which they should assume the worst. But what is unclear is whether such individuals will retrospectively resent the privileged information holders in the same way as those who know they are being kept in the dark may grow frustrated with those who

dodge efforts to gain access to known unknowns. Presumably this is a question regarding the perceived intentionality of the non-disclosure (Ames & Fiske, 2013; Cushman & Young, 2011). One speculation is that unknown unknowns—compared to known unknowns—are less likely to be seen as intentional withholdings. That is, those not thinking about a dimension (on which they actually are ignorant) may mistakenly project this lack of concern onto the privileged information holder, encouraging the thought that they may not have been thinking about the dimension as well. Regardless, this highlights that when one initially does not have access to privileged information, one's evaluations of a target are a function not merely of what one knows but of one's inferences about what one doesn't know (and about why one does or did not know it).

Conclusion

Successfully navigating social and economic life requires people to wisely consider not just the information that others *currently* have, but the information that others might *eventually* have. As we have argued and illustrated, firms, consumers, employees, and friends all regularly navigate this gap between what others know and what others might come to know. We have demonstrated one route through which such better-informed parties err in making such SEEs: Privileged information holders overestimate the extent to which privileged information (vs. information that is common knowledge) tilts others' judgments. Thus, in contrast with work that has shown how privileged information holders do not always accurately navigate *open* information asymmetries (e.g., Camerer et al., 1989), we demonstrate one way which privileged information holders do not accurately navigate potentially *closing* information asymmetries. Decisions to disclose (or withhold) are—at their core—decisions that are informed by how others are estimated to respond upon learning new, previously privileged information. And

insofar as privileged information holders lean on their SEEs about how others would respond if they also knew privileged information, decisions to close information asymmetries—and reveal, for example, sensitive foibles or fulfilling triumphs—may not have the impact one imagines.

References

- Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256–274.
- Ames, D. L., & Fiske, S. T. (2013). Intentional Harms Are Worse, Even When They’re Not. *Psychological Science*, *24*(9), 1755–1762.
- Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290.
- Bazerman, M. H., & Carroll, J. S. (1987). Negotiator cognition. *Research in Organizational Behavior*, *9*, 247–288.
- Becker, J. U., Clement, M., & Nöth, M. (2016). Start-ups, incumbents, and the effects of takeover competition. *Journal of Business Research*, *69*(12), 5925–5933.
- Blau, P. M. (1964). Justice in Social Exchange. *Sociological Inquiry*, *34*(2), 193–206.
- Bond, C. F., Horn, E. M., & Kenny, D. A. (1997). A model for triadic relations. *Psychological Methods*, *2*(1), 79–94.
- Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*(6), 906–921.
- Chambers, J. R., Epley, N., Savitsky, K., & Windschitl, P. D. (2008). Knowing Too Much: Using Private Knowledge to Predict How One Is Viewed by Others. *Psychological Science*, *19*(6), 542–548.

- Cooney, G., Boothby, E. J., & Lee, M. (2021). The thought gap after conversation: Underestimating the frequency of others' thoughts about us. *Journal of Experimental Psychology: General*. Advance online publication.
- Critcher, C. R., & Dunning, D. (2015). Self-Affirmations Provide a Broader Perspective on Self-Threat. *Personality and Social Psychology Bulletin, 41*(1), 3–18.
- Cushman, F., & Young, L. (2011). Patterns of Moral Judgment Derive From Nonmoral Psychological Representations. *Cognitive Science, 35*(6), 1052–1075.
- Dawes, R. M. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology, 25*(1), 1–17.
- Desai, M. N., Lockett, A., & Paton, D. (2018). Information Asymmetries in the Hiring Process and the Risk of New Leader Dismissal: Insights from English Premier League Soccer Organizations. *British Journal of Management, 29*(1), 26–42.
- Dietvorst, B. J., & Simonsohn, U. (2019). Intentionally “biased”: People purposely use to-be-ignored information, but can be persuaded not to. *Journal of Experimental Psychology: General, 148*(7), 1228–1238.
- Donnelly, K., Moon, A., & Critcher, C. R. (2022). Do people know how others view them? Two approaches for identifying the accuracy of metaperceptions. *Current Opinion in Psychology, 43*, 119–124.
- Dunbar, R. I. M. (2004). Gossip in Evolutionary Perspective. *Review of General Psychology, 8*(2), 100–110.
- Dunbar, R. I. M., Marriott, A., & Duncan, N. D. C. (1997). Human conversational behavior. *Human Nature, 8*(3), 231–246.

- Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288–299.
- Friedberg, L., & Stern, S. (2014). Marriage, Divorce, and Asymmetric information. *International Economic Review*, 55(4), 1155–1199.
- Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science*, 1(1), 107–143.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46(2), 107–119.
- Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, 75(2), 332–346.
- Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1–3), 405–440.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. *Social Psychology: Handbook of Basic Principles.*, 133–168.
- Hoch, S. J. (1987). Perceived consensus and predictive accuracy: The pros and cons of projection. *Journal of Personality and Social Psychology*, 53(2), 221–234.
- John, L. K., Barasz, K., & Norton, M. I. (2016). Hiding personal information reveals the worst. *Proceedings of the National Academy of Sciences*, 113(4), 954–959.
- Jones, E. E., Carter-Sowell, A. R., Kelly, J. R., & Williams, K. D. (2009). 'I'm Out of the Loop': Ostracism Through Information Exclusion. *Group Processes & Intergroup Relations*, 12(2), 157–174.

- Jones, E. E., & Goethals, G. R. (1987). Order effects in impression formation: Attribution context and the nature of the entity. *Attribution: Perceiving the Causes of Behavior*, 27–46.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136–153.
- Kenny, D. A. (1994). *Interpersonal Perception: A Social Relations Analysis*. Guilford Press.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin*, 102(3), 390–402.
- Kenny, D. A., Bond, C. F., Mohr, C. D., & Horn, E. M. (1996). Do we know how much people like one another? *Journal of Personality and Social Psychology*, 71(5), 928–936.
- Kenny, D. A., & DePaulo, B. M. (1993). Do people know how others view them? An empirical and theoretical account. *Psychological Bulletin*, 114(1), 145–161.
- Keysar, B., Ginzler, L. E., & Bazerman, M. H. (1995). States of Affairs and States of Mind: The Effect of Knowledge of Beliefs. *Organizational Behavior and Human Decision Processes*, 64(3), 283–293.
- Laing, R. D., Phillipson, H., & Lee, A. R. (1966). *Interpersonal perception: A theory and a method of research*. Springer.
- Lane, J. D., & Wegner, D. M. (1995). The cognitive consequences of secrecy. *Journal of Personality and Social Psychology*, 69(2), 237–253.
- Lewis, G. (2011). Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors. *American Economic Review*, 101(4), 1535–1546.
- Lyubomirsky, S., Tucker, K. L., Caldwell, N. D., Berg, K. (1999). Why ruminators are poor problem solvers: Clues from the phenomenology of dysphoric rumination. *Journal of Personality and Social Psychology*, 77(5), 1041-1060.

- Maaravi, Y., & Levy, A. (2017). When your anchor sinks your boat: Information asymmetry in distributive negotiations and the disadvantage of making the first offer. *Judgment and Decision Making, 12*(5), 420–429.
- Maas, J., Wismeijer, A. A. J., Van Assen, M. A. L. M., & Aquarius, A. E. A. M. (2012). Is it bad to have secrets? Cognitive preoccupation as a toxic element of secrecy. *International Journal of Clinical and Health Psychology, 12*(1), 23–37.
- McKone, E., Davies, A. A., Fernando, D., Aalders, R., Leung, H., Wickramariyaratne, T., & Platow, M. J. (2010). Asia has the global advantage: Race and visual attention. *Vision Research, 50*(16), 1540-1549.
- Mishra, D. P., Heide, J. B., & Cort, S. G. (1998). Information Asymmetry and Levels of Agency Relationships. *Journal of Marketing Research, 35*(3), 277–295.
- Moon, A., Gan, M., & Critcher, C. R. (2020). The overblown implications effect. *Journal of Personality and Social Psychology, 118*(4), 720–742.
- Morewedge, C. K., Giblin, C. E., & Norton, M. I. (2014). The (perceived) meaning of spontaneous thoughts. *Journal of Experimental Psychology: General, 143*(4), 1742–1754.
- O'Brien, E., & Ellsworth, P. C. (2012). Saving the Last for Best: A Positivity Bias for End Experiences. *Psychological Science, 23*(2), 163–165.
- O'Brien, E., Kristal, A. C., Ellsworth, P. C., & Schwarz, N. (2018). (Mis)imagining the good life and the bad life: Envy and pity as a function of the focusing illusion. *Journal of Experimental Social Psychology, 75*, 41–53.
- Parducci, A. (1964). Sequential effects in judgment. *Psychological Bulletin, 61*(3), 163–167.

- Quinn, D. M., & Chaudoir, S. R. (2009). Living with a concealable stigmatized identity: The impact of anticipated stigma, centrality, salience, and cultural stigma on psychological distress and health. *Journal of Personality and Social Psychology, 97*(4), 634–651.
- Quinn, D. M., & Earnshaw, V. A. (2013). Concealable Stigmatized Identities and Psychological Well-Being: Concealable Stigmatized Identities. *Social and Personality Psychology Compass, 7*(1), 40–51.
- Roese, N. J., & Olson, J. M. (1997). Counterfactual Thinking: The Intersection of Affect and Function. In *Advances in Experimental Social Psychology* (Vol. 29, pp. 1–59). Elsevier.
- Rogers, T., & Norton, M. I. (2011). The artful dodger: Answering the wrong question the right way. *Journal of Experimental Psychology: Applied, 17*(2), 139–147.
- Rosenzweig, E., & Critcher, C. R. (2014). Decomposing forecasting: The salience-assessment-weighting (SAW) model. *Current Directions in Psychological Science, 23*, 368–373.
- Samuelson, W. F., Bazerman, M. H., 1985. *Negotiation under the winner's curse*. In: V. Smith (Ed.), *Research in Experimental Economics*, Vol. III, pp. 38-105, Greenwich, CT, JAI Press.
- Savitsky, K., Epley, N., & Gilovich, T. (2001). Do others judge us as harshly as we think? Overestimating the impact of our failures, shortcomings, and mishaps. *Journal of Personality and Social Psychology, 81*(1), 44–56.
- Schulze, W. S., Lubatkin, M. H., Dino, R. N., & Buchholtz, A. K. (2001). Agency Relationships in Family Firms: Theory and Evidence. *Organization Science, 12*(2), 99–116.
- Schwarz, N., Kahneman, D., & Xu, J. (2009). *Global and episodic reports of hedonic experience*. In R. Belli, D. Alwin, & F. Stafford (eds.), *Using calendar and diary methods in life events research* (pp. 157-174). Newbury Park, CA: Sage.

- Shedler, J., & Manis, M. (1986). Can the availability heuristic explain vividness effects? *Journal of Personality and Social Psychology*, *51*(1), 26–36.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013). *Life after P-Hacking*. SSRN.
- Slepian, M. L. (2021). A process model of having and keeping secrets. *Psychological Review*. Advance online publication.
- Slepian, M. L., Chun, J. S., & Mason, M. F. (2017). The experience of secrecy. *Journal of Personality and Social Psychology*, *113*(1), 1–33.
- Slepian, M. L., & Greenaway, K. H. (2018). The benefits and burdens of keeping others' secrets. *Journal of Experimental Social Psychology*, *78*, 220–232.
- Slepian, M. L., Greenaway, K. H., & Masicampo, E. J. (2020). Thinking Through Secrets: Rethinking the Role of Thought Suppression in Secrecy. *Personality and Social Psychology Bulletin*, *46*(10), 1411–1427.
- Slepian, M. L., & Kirby, J. N. (2018). To Whom Do We Confide Our Secrets? *Personality and Social Psychology Bulletin*, *44*(7), 1008–1023.
- Slepian, M. L., Kirby, J. N., & Kalokerinos, E. K. (2020). Shame, guilt, and secrets on the mind. *Emotion*, *20*(2), 323–328.
- Stanley, S. M., Rhoades, G. K., Scott, S. B., Kelmer, G., Markman, H. J., & Fincham, F. D. (2017). Asymmetrically committed relationships. *Journal of Social and Personal Relationships*, *34*(8), 1241–1259.
- Stigler, G. J. (1961). The Economics of Information. *Journal of Political Economy*, *69*(3), 213–225.
- Stiglitz, J. E. (2002). Information and the Change in the Paradigm in Economics. *American Economic Review*, *92*(3), 460–501.

- Tesser, A., & Paulhus, D. L. (1976). Toward a Causal Model of Love. *Journal of Personality and Social Psychology*, 34(6), 1095-1105.
- Thaler, R. H. (1988). Anomalies: The Winner's Curse. *Journal of Economic Perspectives*, 2(1), 191-202.
- Tirole, J. (2015). Market Failures and Public Policy. *American Economic Review*, 105(6), 1665-1682.
- Travers, M., Van Boven, L., & Judd, C. (2014). The Secrecy Heuristic: Inferring Quality from Secrecy in Foreign Policy Contexts. *Political Psychology*, 35(1), 97-111.
- Varnum, M. E. W., Grossmann, I., Kitayama, S., & Nisbett, R. E. (2010). The Origin of Cultural Differences in Cognition: Evidence for the Social Orientation Hypothesis. *Current Directions in Psychological Science*, 19(1), 9-13.
- Wegner, D. M., Lane, J. D., & Dimitri, S. (1994). The allure of secret relationships. *Journal of Personality and Social Psychology*, 66(2), 287-300.
- Welker, M. (1995). Disclosure Policy, Information Asymmetry, and Liquidity in Equity Markets. *Contemporary Accounting Research*, 11(2), 801-827.
- Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 78(5), 821-836.
- Yamagishi, T., Terai, S., Kiyonari, T., Mifune, N., & Kanazawa, S. (2007). The Social Exchange Heuristic: Managing Errors in Social Exchange. *Rationality and Society*, 19(3), 259-291.

Supplemental Materials

Study 1

Results without exclusions. First, we report the results of analyses including the fifty-one participants who failed the attention check question. As in the main text, participants in the privileged information condition believed that the yoked other would evaluate Cabin less positively ($M = 16.46$, $SD = 23.54$) than participants in the non-privileged information condition ($M = 28.75$, $SD = 14.42$), $t(325.22) = 6.28$, $p < .001$, $d = 0.63$.

Results of the exploratory willingness-to-pay measure. Next, we analyzed a measure that we preregistered as exploratory that asked participants to estimate how much the yoked other would be willing to pay for a ~400-mile trip with Cabin (i.e., a trip from Los Angeles to San Francisco). We preregistered the measure as exploratory because we expected great variability in how people would estimate the trip would be monetarily valued, and thus we weren't sure the measure would be appropriate for testing our effect of interest. Additionally, our preregistration specified that we would add \$1.00 to each participant's stated willingness-to-pay and take the natural log of that value. This transformed WTP is what was submitted to the analyses reported below. Participants in the privileged information condition estimated that the yoked other would be willing to pay marginally less ($M_{transWTP} = \$5.32$, $SD_{transWTP} = \$1.69$) than participants in the non-privileged information condition ($M_{transWTP} = \$5.57$, $SD_{transWTP} = \$1.23$), $t(397) = 1.74$, $p = .083$, $d = 0.17$).¹

Study 2: Results Without Exclusions

Next, we report the results of Study 2 including the seventy-five participants who failed the memory-based attention check. Including these participants resulted in a nearly identical

¹ This analysis excluded two participants: One participant who entered a '\$' instead of stating a willingness-to-pay, and one participant who indicated a range (\$500-\$1,000).

result to that which we report in the main text. Participants in the privileged information condition believed that the yoked other would evaluate the Oculus Go less positively ($M = 26.37$, $SD = 18.02$) than did participants in the non-privileged information condition ($M = 31.15$, $SD = 13.01$), $t(297.16) = 2.79$, $p = .006$, $d = 0.31$.

Study 3

Pretest. We recruited one hundred six American AMT workers. These workers evaluated twenty-one different attributes that related to an investment opportunity. All evaluations were completed in a random order; participants indicated their evaluations on a 101-point slider scale anchored at 0 (*much less likely to invest in the stock*) and 100 (*much more likely to invest in the stock*). We identified two features that were relatively neutral, one feature that was relatively positive, and one feature that was relatively negative. The two neutral features were “The company has a massive manufacturing plant located in China” ($M = 49.98$, $SD = 24.33$) and “Several analysts have encouraged current owners of the stock to “hold” the stock – that they should neither buy nor sell the stock” ($M = 49.25$, $SD = 19.08$). The negative feature was “A mid-level executive at the company just quit following a sexual harassment lawsuit” ($M = 30.28$, $SD = 23.82$). The positive feature was “There have been rumors that Google may be interested in acquiring the company” ($M = 70.90$, $SD = 17.26$).

Results without exclusions. Next, we report the results of Study 3 including the eighty-two participants who failed the memory-based attention check. These results are similar to those reported in the main text. First, we examined the SEEs of those in the other perspective condition. Participants estimated that others’ evaluations would be more positive if they did not currently know the fourth feature (privileged information condition: $M = 68.91$, $SD = 18.71$) than if that information was never withheld from them (non-privileged information condition: $M =$

49.13, $SD = 22.68$), $t(457) = 10.23$, $p < .001$, $d = 0.96$. This replicates the PIE. We submitted the evaluation composite to a two-way 2(information: privileged or non-privileged) X 2(perspective: self or social) ANOVA. An Information X Perspective interaction emerged, $F(1, 911) = 16.77$, $p < .001$, $\eta_p^2 = .018$. Further, there was *some* directional truth behind the PIE: Participants in the privileged self perspective condition evaluated the investment opportunity more positively ($M = 56.81$, $SD = 24.26$) than participants in the non-privileged self perspective condition did ($M = 48.99$, $SD = 22.47$), $t(454) = 3.57$, $p < .001$, $d = 0.34$.

Study 4: Results Without Exclusions

In the analyses below, we report the results including the twenty-three participants who failed the attention check question. First, we tested for the PIE. Participants in the privileged/other-judgment condition thought that the yoked other would evaluate the job candidate more negatively ($M = 42.44$, $SD = 22.24$) than participants in the non-privileged information condition ($M = 50.89$, $SD = 23.75$), $t(235) = 2.81$, $p = .005$, $d = 0.37$ (This replicates the PIE observed in Studies 1-3.) Next, we considered whether participants themselves offered more negative evaluations of the target when they held (negative) privileged information—as opposed to information that was common knowledge—about the target. This was not the case: Participants' self-judgments were not more negative (and were actually non-significantly more positive) when they held privileged information ($M = 49.05$, $SD = 24.24$) compared to when such information was non-privileged ($M = 46.75$, $SD = 25.80$), $t < 1$, $d = 0.09$.

Study 5

Pretest. We recruited seventy-seven American AMT workers to evaluate twenty-one attributes of a house. All evaluations were completed in a random order and made on a 101-point slider scale anchored at 0 (*much less likely to recommend the house to a friend*) and 100 (*much*

more likely to recommend the house to a friend). We identified two features that were relatively neutral, one feature that was relatively positive, and one feature that was relatively negative. The two neutral features were “The house has a very nice oak tree in front of it” ($M = 60.97$, $SD = 20.64$) and “The living room has an area that would be perfect for two large bookshelves” ($M = 60.40$, $SD = 22.63$). The negative feature was “The kitchen has a shortage of cabinet and storage space” ($M = 41.05$, $SD = 30.59$). And the positive feature was “Solar panels on the roof keep the electricity bill low” ($M = 80.79$, $SD = 15.54$).

Results without exclusions. In the analyses below, we report the results including the thirty participants who failed the attention check question. These results are nearly identical to those reported in the main text. To begin, we considered those judgments about a target who either actually (non-privileged condition) or potentially (privileged condition) had complete information about the house. We submitted these SEEs to a two-way 2(information: privileged or non-privileged) \times 2(key feature: positive or negative) ANOVA. Critically, we observed a significant Information \times Key Feature interaction, $F(1, 491) = 13.82$, $p < .001$, $\eta_p^2 = .027$. Although the significant interaction reflects a replication of the PIE, we conducted pairwise comparisons to see whether the PIE emerged regardless of whether the key feature was positive or negative.

When the key feature was negative, privileged information participants thought that the yoked other would ultimately evaluate the job candidate more negatively ($M = 63.92$, $SD = 18.87$) than did non-privileged participants who did not think this feature was privileged ($M = 69.81$, $SD = 17.34$), $t(237) = 2.50$, $p = .013$, $d = 0.32$. Mirroring this result, when the self possessed positive information about the job candidate that the yoked other did not, these privileged information holders thought the other would ultimately evaluate the candidate more

positively ($M = 74.34$, $SD = 14.81$) than did those non-privileged information participants who thought the feature was already common knowledge ($M = 68.62$, $SD = 17.97$), $t(247.62) = 2.78$, $p = .006$, $d = 0.35$.

Next, we examined SEEs about targets who knew only three of the four features of the house, either because the other target was said to have actually learned three of the four features (privileged information condition) or because participants had counterfactualized that this knowledge was never actually learned (non-privileged information condition). We submitted these SEEs to a two-way 2(information: privileged or non-privileged) X 2(key feature: positive or negative) ANOVA. As something of a manipulation check, we observed a large main effect of valence, $F(1, 491) = 133.13$, $p < .001$, $\eta_p^2 = .213$. That is, SEEs were more positive [negative] when the other supposedly lacked knowledge of the negative [positive] feature. But crucially, we did not observe an Information X Key Feature interaction, $F < 1$.

Study 6

Pretest. One hundred two American AMT workers evaluated ten features of Move, the online grocery app. Evaluations were completed in a random order, and made on a 101-point slider scale anchored at 0 (*much less likely to use the app*) and 100 (*much more likely to use the app*). We identified three features that were positive and one feature that was negative. The three positive features were “For every order, free 2-day nationwide shipping is guaranteed” ($M = 79.87$, $SD = 16.72$), “For customer service, consumers can chat with a real person” ($M = 75.51$, $SD = 19.44$), and “Shoppers can see the region of the world where each product is produced” (M

= 69.22, $SD = 19.17$). The negative feature was “The membership fee is \$195/year” ($M = 29.51$, $SD = 26.84$).²

Results without exclusions. Next, we report the results from Study 6 that include the forty-two participants who failed the attention check question. These results were similar to those reported in the main text. To begin, we tested whether participants’ SEEs differed between the three information conditions. And indeed, this was the case, $F(2, 1030) = 22.77$, $p < .001$, $\eta_p^2 = .042$. To begin, we replicated the PIE: Participants in the privileged information condition believed that the yoked other would evaluate Move more negatively ($M = 3.45$, $SD = 25.62$) than did participants in the non-privileged information condition ($M = 15.35$, $SD = 20.13$; $p < .001$), $t(1030) = 6.75$, $p < .001$, $d = 0.52$. Second, we found that participants encouraged to desegregate the initially privileged piece of information estimated that the yoked other would be more positive about Move ($M = 9.61$, $SD = 22.83$) than those in the standard privileged information condition, $t(1030) = 3.52$, $p < .001$, $d = 0.25$. And third, we saw that the desegregation intervention eliminated some, but not all of the PIE, $t(1030) = 3.31$, $p < .001$, $d = 0.27$.

Study 7: Results Without Exclusions

In the analyses below, we report the results from Study 7 that include the forty-one participants who failed the attention check question. These results are similar to those reported in the main text. We began by testing for evidence of the privileged information effect. As hypothesized, close others were estimated to evaluate the dating prospect more negatively in the privileged condition ($M = 68.50$, $SD = 19.17$) compared to the non-privileged condition ($M = 74.32$, $SD = 14.65$), $t(229) = 2.56$, $p = .011$, $d = 0.34$.

² In the pretest, the wording used for this negative feature was slightly different. It was “Only members get access to everything that Move sells, and the membership fee is \$195/year.” Similarly, the wording for the third positive item in the pretest was “Consumers can see the region of the world where each product is sourced from.”

Were privileged and non-privileged participants' SEEs equally sensitive to the target feature? Or instead—consistent with the enhanced weighting account—were privileged SEEs (compared to non-privileged SEEs) more closely tied to evaluations of the target feature? We tested this by regressing the general evaluation composite on participants' information condition (-1 = non-privileged, +1 = privileged), the target feature evaluation (standardized), the alternative features evaluation (standardized), as well as the interaction between each feature evaluation composite and the information condition.

Replicating the existence of the PIE, this revealed a main effect of information condition, $\beta = -0.13$, $t(225) = 2.76$, $p = .006$. Further, we also observed main effects of the target feature evaluation, $\beta = 0.35$, $t(225) = 7.58$, $p < .001$, and the alternative features evaluation, $\beta = 0.55$, $t(225) = 11.70$, $p < .001$. Both indicate that participants' general evaluations were sensitive to how they expected their close other to evaluate the dating prospect's individual features. We next turn to the interaction terms. There was no Information Condition X Alternative Features Evaluation interaction, $\beta = -0.02$, $t < 1$. We did, however, observe a significant Information Condition X Target Feature Evaluation interaction, $\beta = 0.14$, $t(225) = 2.93$, $p = .004$. When the unflattering fact about the dating prospect was privileged, evaluations of this feature strongly informed participants' estimates, $\beta = 0.49$, $t(225) = 7.36$, $p < .001$; when the unflattering feature was common knowledge, this effect was reduced, $\beta = 0.22$, $t(225) = 3.32$, $p = .001$.

Supplemental Study A

Study 3 found that information learned in a second stage had a disproportionate influence on the self's own evaluations of a target. The goal of Supplemental Study A was to test the robustness of this pattern. To that end, participants in Supplemental Study A were assigned either to have complete information about a target or to learn (previously withheld, and thus privileged)

information about a target in a second stage. Participants then provided their own evaluations of the target (i.e., an investment opportunity). These conditions largely mirror the two *self-perspective* conditions from Study 3.

However, in Supplemental Study A, privileged information condition participants—those who had information initially withheld from them—were not asked to provide an *initial* evaluation of the target based on the first 3 features they learned. Recall that, in Study 3, participants in the self-perspective privileged condition offered this evaluation before learning the fourth feature, at which point they provided a new evaluation. Although we found that the PIE itself emerges regardless of whether participants were explicitly asked to provide an SEE of the target based on what they presently know before offering an SEE about how the other would feel with complete information (see Studies 1 and 4 for studies that omit this feature), we were curious whether the significant (but smaller-than-estimated) effect of privileged information on participants' *own* evaluations would be similarly robust to this change.

Participants and design. Four hundred eighteen Americans were recruited from AMT in exchange for nominal compensation. Participants were randomly assigned to one of two information conditions: *privileged* or *non-privileged*.

Materials and procedure. The materials and procedure were largely identical to those in the two self-perspective conditions in Study 3. All participants initially read that they would be assisting a venture capital company (Heritage Capital) forecast stock performance. Participants learned that they would be (ostensibly) yoked with a future participant, and that all participants would receive information about different features of the investment opportunity. As in Study 3, a graphic with four cards—each representing one of the four features of the investment opportunity—was displayed to all participants.

Information manipulation. Participants' information condition determined the amount of information that they themselves received. Participants in the *non-privileged* information condition learned that both they themselves and the yoked other would receive all four pieces of information about the investment opportunity. Participants in the *privileged* information condition learned that one person—either they themselves or the yoked other—would receive three pieces of information, while the other would receive all four.

All participants then saw three of the investment opportunity's features—one negative feature and two neutral features—in a random order. Then, participants in the non-privileged information condition—and the ostensible yoked other—learned about the fourth (positive) feature of the investment opportunity. These participants then provided their own evaluations of the investment opportunity.

Conversely, participants in the privileged information condition learned that the yoked other—but not they themselves—was learning this fourth feature. Then, these participants were asked to *consider* how they themselves—with only these three pieces of information—would evaluate the investment opportunity. That is, these participants did not provide a rating on a slider scale; they were merely asked to consider how they would evaluate the investment opportunity. Next, the information asymmetry was resolved: Participants in the privileged information condition actually did receive the investment opportunity's fourth feature. Finally, having received complete information, these participants evaluated the investment opportunity.

The evaluation composite. All participants completed two items. The first item asked participants how they would evaluate the investment opportunity. This was made on a 101-point slider scale anchored at -50 (*very negatively*) and +50 (*very positively*). The second item asked participants how likely they would be to invest their own money in Cedar Technologies, Inc.

This measure was completed on a 101-point scale anchored at 0% and 100%. These items were placed on equivalent 0-100 scales and then averaged to form an *evaluation* composite ($r = .76$).

Following these general evaluations, all participants evaluated each individual feature of the investment opportunity. The four features were evaluated one at a time, in a random order. Each judgment was made on a 101-point slider scale. These scales were anchored at -50 (*very negatively*) and +50 (*very positively*). Before any analyses were conducted, these items were transformed to be on 0-to-100 scales.

Results and discussion. Did participants who had positive information initially withheld from them ultimately develop more positive evaluations of the investment opportunity as compared to those who had learned this information from the beginning? We found no evidence that this was the case. That is, there was not a significant difference between how participants in the non-privileged information condition ($M = 48.04, SD = 23.30$) and the privileged information condition ($M = 49.39, SD = 24.28$) evaluated the investment opportunity, $t < 1$. Whereas in Study 3 we found that learning this information in two stages led it to be more influential in participants' own evaluations than if that (sometimes privileged) feature had simply been learned with the other information, the present study was unable to replicate these results.

We proceeded to test whether participants' own evaluations of the individual features differed by information condition. To do this, we defined two variables. One was the *target features evaluation*. This reflected participants' evaluations of the positive feature (which for privileged participants, was initially hidden from them). The second was the *alternative features evaluation*. This reflected an average of the other three evaluations of the features (that were always common knowledge). There was not a significant difference in how participants in the non-privileged and privileged information conditions evaluated the target features ($M_{non-privileged} =$

76.64, $SD_{non-privileged} = 21.13$, $M_{privileged} = 75.11$, $SD_{privileged} = 20.15$, $t < 1$) or alternative features ($M_{non-privileged} = 46.70$, $SD_{non-privileged} = 18.36$, $M_{privileged} = 45.49$, $SD_{privileged} = 20.98$, $t < 1$).³

Next, we tested whether—across the privileged versus non-privileged information conditions—participants’ own evaluations of the individual features possessed different predictive power of their overall evaluations of the investment opportunity. We tested this by regressing the evaluation composite on participants’ information condition (-1 = non-privileged, +1 = privileged), the target feature evaluation (standardized), the alternative features evaluation (standardized), as well as the interaction between each feature evaluation composite and the information condition. Crucially, there was not a significant Information X Target Features Evaluation interaction, $\beta = .02$, $t < 1$. Furthermore, there was no Information X Alternative Features either, $\beta = -.07$, $t(412) = -1.88$, $p = .061$. In other words, participants in the privileged and non-privileged conditions did not differ in the extent to which they weighted the target feature in their own evaluations.

In sum, Supplemental Study A found no evidence that initially withheld information had an outsized impact on participants’ *own* evaluations of a target. This (null) result is in contrast to Study 3, which found that learning information in a second stage did sway evaluations (although, not as much as privileged information holders predicted). One possibility is that a methodological difference between Study 3 and Supplemental Study A explains this difference. Participants in the present study did not initially evaluate the target (based on the three attributes they first learned) before offering their final evaluation (based on all four attributes). It is possible that making the first evaluation encouraged (privileged information) participants to

³ Examining participants’ evaluations of each of the four features individually does not change this conclusion, all $ts < 1.60$, all $ps > .111$.

incorporate the final piece of information in a more segregated manner. That said, note that the PIE itself—which applies to SEEs, not the self’s own evaluations—does not depend on this feature. We consistently observed evidence of the PIE regardless of whether those making SEEs first did (e.g., Studies 2-3) or did not (e.g., Study 1, 4) offer an initial SEE before a final SEE about a target with complete information. Note how these patterns are easily accommodated by the segregated processing account. Privileged information holders naturally differentiate two SEEs—one based on what someone else does know, and one based on what someone else *could* know—even when not prompted. Information acquisition to inform one’s own opinions almost always occurs through a sequential process (i.e., learning one fact before another, before another). In turn, people’s own evaluations may not be sensitive to the sort of segregation processes that are core to the PIE (and thus more easily emerge with SEEs) unless people are prompted to explicitly form an evaluation before learning more information.

Supplemental Study B

By our proposed segregated processing account, privileged information exerts an undue influence on SEEs because it is not incorporated as one among many pieces of information that define another’s representation of a target. In many of our studies, we asked participants how another would respond if they “did” learn the privileged information. Note that there is an ambiguity of interpretation here, but one that should not be critical to our theoretical logic. This could reference a counterfactual *past* possibility (“if instead of only the limited information they learned, they did learn this as well”) or a subjunctive *future* possibility (“if they were to learn this information in the future”). By our logic, either interpretation should produce the SEE, because each still requires participants to incorporate new information into an SEE in a segregated way.

Supplemental Study B tested whether the PIE actually does emerge (and emerge relatively similarly) even when the materials more clearly invite one interpretation or the other.

Participants and design. Nine hundred forty-six Americans were recruited from AMT in exchange for nominal compensation. Participants were randomly assigned to one of three information conditions: *privileged-future*, *privileged-past*, and *non-privileged*. One hundred forty-two participants failed a memory-based attention check and were excluded from all analyses reported below. This left a final sample of 804 participants.

Materials and procedure. The materials were largely identical to those utilized in Study 3. All participants initially read that they would be assisting a venture capital company (Heritage Capital) forecast stock performance. Participants learned that they would be (ostensibly) yoked with another participant, and that all participants would receive information about different features of the investment opportunity. A graphic with four cards—each representing one of the four features of the investment opportunity—was displayed to all participants.

Information manipulation. To begin, participants learned what they and the yoked other would learn about the target investment opportunity. Participants in the *non-privileged* information condition learned that both they themselves and the yoked other would receive all four pieces of information about the investment opportunity. Participants in both the *privileged-future* and *privileged-past* information conditions learned that one person—either they themselves or the yoked other—would receive three pieces of information. The other person—i.e., the one who did not receive only three pieces of information—would receive all four pieces of information.

Three pieces of information about the investment opportunity—one negative and two neutral—were revealed one-by-one to participants in a random order. What happened next varied

by the information condition to which participants were randomly assigned. In the non-privileged information condition, the fourth (positive) piece information was shown to both the participant and (ostensibly) the yoked other. Then, participants in the non-privileged information condition offered estimates of how the yoked other—who knew all four features—would evaluate the investment opportunity.

In the privileged-future and privileged-past information conditions, only participants themselves initially learned about this fourth, positive feature. That is, the information was (ostensibly) hidden from the yoked other. After this fourth and final feature was revealed, participants in both the privileged-future and privileged-past conditions estimated how the yoked other—who was said to know only three of the features—would evaluate the investment opportunity. After these participants provided these initial SEEs, they considered that the information asymmetry would be resolved in the future (*privileged-future*) or considered that the information asymmetry was resolved in the past (*privileged-past*). More specifically, participants in the privileged-future information condition were asked, “Imagine that the other participant were to learn about the investment opportunity’s Feature C.” In contrast, participants in the privileged-past information condition read, “The other participant ultimately did learn about the investment opportunity’s Feature C.” Participants in both privileged information conditions then provided estimates of how the other—if they were to learn or did learn about the fourth, positive feature—would evaluate the investment opportunity.

The evaluation composite. All participants completed two items. The first asked participants to estimate how the yoked other would evaluate the investment opportunity overall. This was made on a 101-point slider scale anchored at -50 (*very negatively*) and +50 (*very positively*). The second item asked participants to indicate how likely the yoked other would be

to invest their own money in Cedar Technologies, Inc. This measure was completed on a 101-point slider scale anchored at 0% and 100%. These items were placed on equivalent 0-to-100 scales and then averaged to form an *evaluation* composite ($r = .73$).

Results and discussion. In all analyses that we report below, we compared participants' SEEs about others who had equivalent information. That is, non-privileged information participants' SEEs were compared against the SEEs that participants in the privileged-future and privileged-past information conditions made about another who had learned in the past or would learn in the future about all four of the features. We first tested if there was any difference in SEEs between the three information conditions. A one-way ANOVA revealed a large omnibus effect of condition, $F(2, 801) = 46.40, p < .001, \eta_p^2 = .104$. We conducted two pairwise comparisons that tested for the PIE in two ways. First, participants' SEEs in the privileged-future information condition were more positive ($M = 67.59, SD = 19.13$) than participants' SEEs in the non-privileged information condition ($M = 52.92, SD = 20.87$), $t(801) = 8.44, p < .001, d = 0.73$. Second, participants' SEEs in the privileged-past information condition ($M = 66.75, SD = 19.54$) were more positive than participants' SEEs in the non-privileged condition, $t(801) = 8.19, p < .001, d = 0.68$. Finally, we found that participants' SEEs did not differ between the privileged-future condition and the privileged-past condition, $t < 1$.

In sum, Supplemental Study B found that the PIE emerged both when an information asymmetry was described as being closed in the future or already closed in the past. That the PIE was roughly equivalent in size between these two conditions reinforces that it is the segregated processing of privileged information—and not an information asymmetry being open or closed—which produces the PIE. Furthermore, it suggests that in our studies for which there was some ambiguity of interpretation—about whether the information asymmetry already had closed

or would ultimately close—the results seemingly did not depend on how participants resolved that ambiguity. This is also reassuring concerning the external validity of the PIE. After all, in reality there will often be ambiguity about whether an information asymmetry already has or will soon close. As long as the initially privileged information is incorporated in a segregated way, the PIE is likely to emerge.