We appreciate the time and thought that Karen Cerulo, Paul DiMaggio, Doug Maynard, and Stephen Vaisey put into critiquing our paper. It is the greatest honor to have one’s work seriously engaged, and this is exactly what our critics have done; for this, we thank them. Taken together, their essays shed greater light on a range of methodological issues that ought to concern every scholar interested in social action. They also help us identify a suite of research strategies that sociologists can use to avoid the attitudinal fallacy. Rather than pursue a point-by-point rebuttal to all four critics, we delve deeper into our argument by concentrating our response on three central points raised in this symposium: (1) While we acknowledge that there are occasions when reported attitudes are correlated with situated behavior, it is our contention that correlations are never high enough to presume equivalence (this presumption is the attitudinal fallacy) and that sociologists have only a rudimentary understanding of the conditions under which they can expect correspondence or divergence; (2) Similarly, sociologists routinely surmise that self-reported behavior is, more or less, an accurate stand-in for actual behavior, glossing the problematic of the fallibility of accounts—we call this the accounting fallacy; and (3) The attitudinal fallacy has been and continues to be a central problem in a variety of social science disciplines, and scholars have converged on the same core solution that we propose in our paper: direct observation of behavior and an appreciation for the situational determinants of action.

* Direct correspondence to Colin Jerolmack, Department of Sociology, 295 Lafayette St., Floor 4, New York, NY 10012; email: jerolmack@nyu.edu. We thank Rodney J. Andrews, Philipp Brandt, Paula England, Alix Rule, and Iddo Tavory for comments and critiques on a prior draft.
WHEN DO ATTITUDES PREDICT BEHAVIORS?

We agree that, “Sometimes, what people think/say actually does inform what they do” (Cerulo 2013: xx).1 “Talk is Cheap” did not prescribe “strong situational determinism”—that is, “the idea that behavioral explanation requires only knowledge about social context and nothing about persons” (Vaisey 2013: xx). The ethnographic commitment to documenting how actors construct a “definition of the situation” is premised on the belief that how people act in the world is contingent on how they interpret it—which encompasses their beliefs, schemas, frames, dispositions, and so on. *Ethnography includes interviewing*. But careful ethnographers interpret verbal data *in relation to observations* of actors’ situated behavior across a variety of social settings. Moreover, they view the interview situation itself as a form of “talk-in-interaction” (Schegloff 1997; Maynard 2003) in which the interviewee’s expressed subjectivity is a contingent social accomplishment. Thus, the kind of ethnography we espouse neither naïvely presumes that verbal data are indicative of stable mental states and behavioral intentions nor discounts verbal data as irrelevant for understanding action.

Cerulo and Vaisey both point to examples of attitude-behavior correlation (ABC) that float around .50, “mirror[ing] the magnitude of other relationships our field accepts with little trepidation—i.e., the association between father-son occupational status” (Cerulo 2013: xx). We are not surprised that there are instances where attitudes and behaviors are related to one another. But we are surprised that a correlation of .50 appears to be offered up as a defense of research

---

1 We concede Cerulo's (2013) point that a number of culture and cognition studies go beyond accounts, and that there is nothing inherent in toolkit theory—or network analysis (Vaisey 2013: xx)—that demands an individualistic approach to social behavior. “Talk is Cheap” simply argued that the “culture in action” models we criticized “explain and analyze social processes strictly at the individual level.” Given that our engagement with this field in the original paper was secondary to the attitudinal fallacy, in this response we focus exclusively on methodological issues.
that treats attitudes and behaviors as equivalent phenomena. If anything, we should conclude the exact opposite. Imagine attempting to publish a paper that operationalizes child’s occupational status by using father’s occupational status. Reviewers would correctly object that the correlation between these two variables is far too low to treat them as if they were the same. Yet while we would not tolerate such presumed equivalence for variables like occupational status, we routinely do for papers that use attitudes to represent behaviors. And often, the situation is worse than that proposed above: as we point out in “Talk is Cheap,” there are several major fields of interest, such as racial attitudes, where ABC is closer to 0 (Pager and Quillian 2005). It is our position, then, that scholars must first ask if there is any empirical evidence for using a particular attitude as a proxy for a behavior. The standard for such evidence must not be that there is a relationship, but rather that the relationship is so strong that one can presume equivalence.

Further, we must be circumspect about what counts as strong evidence. Highly correlated A-B findings are littered with caveats that scholars should be attuned to. For instance, the meta-analysis that Vaisey (2013: xx) points to as proof of a “fairly strong baseline” correspondence between attitudes and behaviors only considers studies in which experimenters created “novel attitudes” (Glasman and Albarracin 2006: 818) about unfamiliar and often “relatively irrelevant” objects (e.g., puzzles or video games) in laboratory settings with college students (ibid.: 807). The meta-analysis did not rely upon “hundreds of studies” but instead 29, defining out of existence the studies that did not meet the authors’ narrow (experimental) parameters. Moreover, almost half the experiments measured behavior immediately after measuring de novo attitudes, and the longest gap between A-B measurement was two weeks. For scholars interested in testing whether durable attitudes or dispositions predict consequential “real-world” behaviors in the general population, such laboratory studies hardly provide adequate validation of ABC.
We are not arguing that attitudes and behaviors are never correlated, but we do contend that it is a fallacy to presume that they are without evidence pertinent to the specific attitude(s) and behavior(s) in question. Importantly, we do not disagree with DiMaggio’s (2013: xx) argument that “contexts selectively reinforce or suppress” cultural repertoires and Cerulo’s (2013: xx) contention that “low to moderate AB correlations…signal a complicated association tempered by intervening factors”—both claims imply the need to study situated action in order to adjudicate between the role of dispositions and situations in shaping behavior (for an exemplar of this approach, see Brubaker at al. 2006).

WHEN CAN WE TRUST SELF-REPORTS OF BEHAVIOR?

A related issue is whether we can trust self-reports of behavior. Vaisey (2013) argues that there is little to worry about, pointing to an article (Kreuter et al. 2008) showing that nine self-reported behaviors gathered through three different survey modes had a mean correlation of .85 with “objective” administrative records. This is exactly the kind of finding we wish we had more of, as it suggests a context within which self-reports strongly correlate with behaviors (although they are still not equivalent). Yet even in this best-case scenario, once again there are important caveats. This study consists of a very circumscribed set of participants (University of Maryland alumni) and a limited set of practices (e.g., whether or not they failed a class or donated money to the college). While the authors conclude from the study that social desirability bias exerted a minimal effect on the accuracy of web-based self-reports, they cannot rule out that the participants may have been motivated to tell the truth because they knew that the investigators—who were sponsored by the UMD alumni office—could independently access their records. Indeed, social desirability bias is a major issue for any research that asks “sensitive questions”
about behavior: respondents seem to underreport the use of illicit drugs and alcohol, smoking, abortion (by as much as half), bankruptcy, and energy consumption; and they have been found to overreport energy conservation, seat belt use, having a library card, church attendance, and exercise (Tourangeau and Yan 2007: 863)—among other things. While DiMaggio (2013: xx) points out that “extensive literatures” describe methods for detecting social desirability bias, Tourangeau and Yan (2007: 863) contend that many of these studies in fact “lack validation data and assume that whichever [verbal] method yields more reports of the sensitive behavior is the more accurate method; survey researchers often refer to this as the ‘more is better’ assumption” (see also Kreuter et al. 2008: 848). We do not wish to say that because of this ambiguity we can never rely upon what people say; instead, our argument is that we need to better understand when we can take accounts as proxies for actions. 2 To simply presume that self-reported behaviors are accurate is to commit an error akin to the attitudinal fallacy—perhaps we could call it the accounting fallacy.

As DiMaggio (2013: xx) mentions, social desirability bias is just one of many reasons why self-reports might “fail to elicit useful indicators of behavior.” Unfortunately, researchers know far less about other reasons. Sociologists have no general theory about when and why verbal data fail to accurately reflect social action, and we are less sanguine than DiMaggio about finding universal solutions. Our skepticism is grounded in over a half-century of interactionist and ethnomethodological studies that challenge a necessary presumption of survey research: that

2 Vaisey’s (2009) own work glosses the dilemma. Although his measures of behavior are verbal responses by teenagers to “sensitive questions,” he does not discuss threats to validity such as social desirability bias. If everyone in Vaisey’s sample was equally likely to underreport “deviant” behavior, or if social desirability bias was randomly distributed, this would not be much of a problem. But we have no way of knowing whether this is the case; and we are unable to rule out the possibility that something like high susceptibility to social desirability bias may be driving responses to both Vaisey’s IV and DV: choosing the moral schema “I do what God tells me” and reporting the lowest levels of deviance. Or, perhaps those who “really” do attempt to do what God tells them feel the most pressure to underreport deviance.
all respondents interpret the question in the same way, and that all of those who choose a particular response mean the same thing (of course, some questions leave far less room for interpretation than others, e.g., “When were you born?”). A common problem in survey, and to a lesser degree interview, research is that standardization pursues reliability at the expense of validity (Cicourel 1964). Studies that examine the survey interview as a form of “talk-in-interaction” reveal the extent to which the “degree of uniformity participants achieve” is “contingent upon the vicissitudes and contingencies of their actions and reactions to one another in and through talk” (Maynard and Schaeffer 2005: 12). While survey researchers often presume that the low number of participants who answer “don’t know” to a question stands as evidence that it is “easy” to answer, Maynard and Schaeffer (2005: 13) detail how survey respondents commonly offer initial answers that depart from the coding scheme. Such responses are washed out of the final results, however, because interviewers regularly work to ensure that they can secure a standardized answer—even if it means departing from the “rules of standardization” to get it. Conversation analysis of survey research reveals respondents routinely hedging their answers (e.g., “I guess”) or offering elaborations, after having given a standardized answer, that resist the implications of the categorization and in some cases appear to conflict with their coded answer. Although some interviewers returned to the original question after such elaborations to seek clarification, “interviewers can and do treat the speech material following a codable answer as so much chaff” (ibid.: 22). In such cases, it is unclear what their coded answer “really” means.

While it remains an empirical question the extent to which “standardization-in-interaction” may impact the validity of survey research, we do not cite this example simply to be contrarian. Importantly, it highlights how any occasion in which verbal data is elicited is an *interaction*, and one in which the respondent (and questioner, if human) constructs “ad hoc”
procedures and meanings in order to make the coding scheme intelligible (Cicourel 1964; Garfinkel 1967; Maynard and Schaeffer 2005). This challenges the presumption that survey questions are a simple extraction of respondents’ pre-formed, decontextualizable categories of thought.

Similar problems can emerge even in what appear to be simple behavioral “counting” exercises. For instance, the survey that Vaisey (2009) relies upon for his behavioral data on “deviance” asks respondents, “In the last year, how often, if ever” they did things they hoped their mother or father would never find out about, cheated on a school assignment, or lied to one of their parents. The choices are: very often; fairly often; sometimes; occasionally; rarely; never; and don’t know. Even presuming that social desirability is not an issue, the answer choices present a major problem for standardization. For one, “the last year” is a very long time horizon from which to recall these events. Given that “respondents use as little cognitive effort as they can to answer survey questions” (Vaisey 2009: 1688), we have some reason to doubt that people would take the time to mentally walk through the last year of their life just to ensure they provide an accurate answer to a single questionnaire item—they might just pick the “response that ‘feels right’” (p. 1689). Perhaps more problematic, we have no idea how each respondent defined categories like “fairly often” and “occasionally.” One person might consider cheating once a month to be “very often;” another might consider cheating once a month to be “rarely.” We should treat these reports like the pain scale that doctors administer to patients: they may provide hints about each individual’s subjective interpretation of their situation, but we cannot objectively compare them as if they represented equal values on an ordinal ranking.

---

3 While most people may be able to accurately recall the last time they bought a house or the first time they had sex, it seems plausible that many would need to keep an incredibly detailed journal to accurately report even the general frequency range (e.g., 5-10) of quotidian practices. Some surveys take steps to resolve this problem by asking more specific questions such as, “Did you consume drugs or alcohol in the past week?” While this may substantially limit the breadth of behavioral data, it likely improves accuracy.
To varying degrees, what Cicourel (2004) refers to as the “ecological validity problem” (treating the meaning of questions as self-evident) can be generalized to a large swath of survey items—interpretation is always involved, and if we have no operating knowledge of members’ meanings we are at a severe disadvantage in figuring out what their responses mean. The implication is that we will never be able to come up with a general solution to the problem of what verbal data tells us about behavior that does not include an accounting of actors’ lifeworlds (Young 2004) and situated social practices. While Maynard (2013) adds an important proviso by pointing out that ethnographers sometimes also commit the error of providing explanations for social action that neglect “the interaction order,” we believe that the “affinities” (Maynard 2003) between ethnography and his preferred method—conversation analysis (CA)—outnumber the dissimilarities when both are compared to verbal methods like interviews and surveys.

**THE SAYING/DOING PROBLEM: BEYOND SOCIOLOGY**

Vaisey (2013: xx) refers to our “resurrection” of the ABC problem as a “misguided debate long since laid to rest.” In one sense he is correct: in sociology’s sister disciplines, this debate has been settled to some extent; but the emerging consensus is the opposite of Vaisey’s—and much of sociology’s—position. Scholars in both economics and psychology have determined, rightly in our view, that they must be far more skeptical when using verbal data to explain behavior. For

---

4 As DiMaggio (2013) points out, interviews have an advantage over surveys here because they can probe for meaning. This is an important point. Long ago, Cicourel (1964) observed the irony that survey researchers routinely rely on methods that lack generalizability (e.g., focus groups) in order to establish the validity of questions from which they generalize the answers to entire populations. Salganik and Levy (2012: 1) develop a novel workaround to this dilemma, the wiki survey, which “combines the quantifiability of a survey and the openness of an interview.”

5 While we are sympathetic to Maynard’s critique, we refute the charge that ethnography haphazardly invokes structure. The kind of ethnography we espouse shares ethnomethodology’s inclination to invoke structure only to the extent that it can be “seen” in situated interaction. The issue boils down to the fact that ethnomethodologists generally have a much more circumscribed view of structure than most sociologists, one that revolves around rules and conventions of interaction (e.g., turn-taking).
example, there is a renewed interest in the “fundamental attribution error” in psychology and marketing research. Smith and Semin (2007: 132) write,

As of a decade or so ago, most researchers would have agreed that symbolic representations such as stereotypes are abstract, stable, and general knowledge structures (or schemas); that they are activated automatically and independent of the perceiver’s goals…and that their activation makes their content available and likely to influence the perceiver’s judgments and actions, even against the perceiver’s wishes. Against this accepted wisdom, they argue that “more recent evidence suggests that stereotypes’ effects on social judgments and social behaviors are extremely malleable and sensitive to details of current social situations” (p. 132). For instance, it has often been noted by psychologists that people “tend to explain other people’s behavior in terms of those people’s inner personality characteristics rather than in terms of the demands of social situations. This tendency has been viewed as automatic, fundamental, and linked to properties of abstract mental processes” (p. 133). However, in a study (Norenzayan and Schwarz 1999) that asked participants to explain the causes for a mass murder they read about in a newspaper article, it was found that their explanations were more situational if the letterhead on the questionnaire read, “Institute for Social Research” and more dispositional if the letterhead read, “Institute of Personality Research.” This demonstrates “the susceptibility of this supposedly fundamental and automatic attribution process to contextual influences” (Smith and Semin 2007: 133).

Smith and Semin (2007) point to numerous studies that indicate that cognition is situational and argue for a “theoretical approach that makes interdependence and mutual constraint between person and context a central focus,” not merely an “unintegrated theoretical ‘add-on’” (p. 134). Schwarz (2006: 19-20) also observes the “high context dependency of attitude judgments,” and he points to over a dozen studies that “present models of judgment and behavior that account for the bulk of the accumulated attitude findings without assuming that people have enduring predispositions.” Indeed, at the very least there is a sizeable movement
against “dispositionalism” in social psychology, one that recommends interactional and situational analyses of social action as the most appropriate method for avoiding the attitudinal fallacy (Haney and Zimbardo 2009; Ross and Nisbett 1991).  

Similarly, areas of economics that have relied upon survey research have been critically reevaluated in light of ABC problems. Indeed, the now-classic notion of “revealed preferences” was, in effect, a methodological admonishment of economists who spend their time asking people about their purchasing habits rather than observing what they actually buy. And in his recent devastating critique of contingent valuation research, Jerry Hausman (2012: 43) concludes that, “respondents to contingent valuation surveys are often not responding out of stable or well-defined preferences, but are essentially inventing their answers on the fly, in a way which makes the resulting data useless for serious analysis.” This leads Hausman (p. 44) to the same basic conclusion that we arrive at: “Put simply, what people say is different than what they do.”

The explosion of field research in economics is in fact a reaction to a general dissatisfaction with verbal methods. “Field experiments,” including but not limited to audit studies, have grown out of a belief that context effects can be so strong that even experiments—with their contrived social settings—are of limited utility in predicting and explaining “real” economic behavior. One of the leading proponents of field experiments points out that they have now “touched nearly every subfield of economics” (List 2011: 7), and Chang et al. (2009: 518) argue that the reason why is because “There is perhaps no more important question for researchers working with survey and experimental preference elicitation methods than whether the elicited values accurately predict real-world behavior.” For instance, List (2006) shows that sports cards dealers in a laboratory setting were “driven strongly” by a preference for “positive

---

6 While Vaisey (2013: xx) claims, “we now know that…the Milgram and Zimbardo experiments never really supported strong situational determinism,” Haney and Zimbardo (2009) offer a powerful defense of a situational interpretation of the Stanford Prison Experiments.
reciprocity,” but that the same sport traders in a natural field experiment “behaved more selfishly” (Levitt and List 2008: 2010). In this and other field experiments, it is found that “The social preferences so routinely observed in the lab”—and, one could add, in measures based on verbal data—“were attenuated in the field” (Levitt and List 2007: 160). Importantly, Levitt and List (2007: 160) add that, “these results are consistent with the wealth of psychological literature that suggests there is only weak evidence of cross-situational consistency of behavior.” They conclude that this is because of one of two possibilities, both of which are compatible with the argument that we have advanced: “(a) there is not a general cross-situational trait called ‘social preferences,’ and/or (b) the subjects view one situation as relevant to social preferences and the other as irrelevant” (p. 160-61). Given this conclusion, studying settings “more representative of the behavior about which economists are seeking to learn” (p. 171) is an integral part of explaining economic action.

CONCLUSION

Sociology has a lot to learn from other disciplines about how to identify, and avoid, the attitudinal and accounting fallacies. Cerulo (2013: xx) is certainly correct that it is not always feasible to directly study the “action sites” we are interested in (e.g., sex); and we agree with her that “our question must be: when can we trust accounts?” While the field of sociology is equipped with the tools to make progress in answering this question, DiMaggio (2013: xx) concurs with our assessment that, “relatively few investigators use the methods available to them.” We can obtain a lot of useful information from verbal data, including people’s categories of thought and self-identifications (Lamont 1992). But if our goal is explaining social action, we must always be wary of the limits of what such data allow us to say about situated behavior. We
cannot simply presuppose that schemas and dispositions enact behaviors in a straightforward way, nor can we presume that if we can come up with clever ways to assess social desirability bias that the self-reports of behavior we gather will be accurate. While ethnography is useful for mapping the saying/doing relationship, it is no panacea. Our discipline is not anywhere close to developing a general theory about the nature of the relationship between verbal data and social action (which, we should be clear, includes “talk-in-interaction”). We hope that researchers of all methodological stripes view this symposium as an invitation to help our discipline work toward such an understanding.

REFERENCES


———. 2013. “The ‘Attitudinal Fallacy’ is a Fallacy: Why We Need Many Methods to Study Culture.” Sociological Methods and Research #(#): xxx-xx.