

# How should we measure Americans' perceptions of socio-economic mobility?

Lawton K. Swan\*    John R. Chambers†    Martin Heesacker\*    Sondre S. Nero‡

## Abstract

Several scholars have suggested that Americans' (distorted) beliefs about the rate of upward social mobility in the United States may affect political judgment and decision-making outcomes. In this article, we consider the psychometric properties of two different questionnaire items that researchers have used to measure these subjective perceptions. Namely, we report the results of a new set of experiments ( $N = 2,167$  U.S. MTurkers) in which we compared the question wording employed by Chambers, Swan and Heesacker (2015) with the question wording employed by Davidai and Gilovich (2015). Each (independent) research team had prompted similar groups of respondents to estimate the percentage of Americans born into the bottom of the income distribution who improved their socio-economic standing by adulthood, yet the two teams reached ostensibly irreconcilable conclusions: that Americans tend to *underestimate* (Chambers et al.) and *overestimate* (Davidai & Gilovich) the true rate of upward social mobility in the U.S. First, we successfully reproduced both contradictory results. Next, we isolated and experimentally manipulated one salient difference between the two questions' response-option formats: asking participants to divide the population into either (a) "thirds" (tertiles) or (b) "20%" segments (quintiles). Inverting this tertile-quintile factor significantly altered both teams' findings, suggesting that these measures are inappropriate (too vulnerable to question-wording and item-formulation artifacts) for use in studies of perceptual (in)accuracy. Finally, we piloted a new question for measuring subjective perceptions of social mobility. We conclude with tentative recommendations for researchers who wish to model the causes and consequences of Americans' mobility-related beliefs.

Keywords: inequality, social mobility, fairness, justice, political ideology, income.

## 1 Introduction

In the January 2015 issue of *Perspectives on Psychological Science*, Davidai and Gilovich claimed to have exposed a startling glitch in the American public's perception of socio-economic reality: despite a well-documented trend of declining faith in the "American Dream" of equal opportunities for upward advancement (Pew Research Center, 2012), most participants in Davidai and Gilovich's investigations ( $N = 3,034$  nationally-representative adults and 290 Mechanical Turk workers) *overestimated* the number of Americans who manage to improve their social class ranking during their lifetime. For instance, when asked to gauge the likelihood that an individual born into the poorest quintile (20%) of the income distribution would remain in the bottom 20% in adulthood, participants' average estimate was 33% — a considerable degree of undue optimism, given that, according to data reported by the Pew Economic Mobility Project (2012), the true rate of immobility for this stratum is closer

to 43%. Thus, even as they acknowledge that the United States falls short of its egalitarian vision, it would seem that most Americans continue to view socio-economic prospects for the poor through rose-tinted spectacles.

Unbeknownst to Davidai and Gilovich, we (Chambers, Swan and Heesacker) had been working in parallel to study the very same phenomenon — we too had attempted to quantify Americans' (in)accurate perceptions of upward social mobility, and we too successfully published our findings (in the journal *Psychological Science*; Chambers, Swan & Heesacker, 2015). The two papers appeared in print within weeks of each other, and both reported a sizable mean-level distortion in respondents' upward mobility appraisals. However, whereas Davidai and Gilovich described a robust trend of overestimation (estimating more mobility than there really is), we found just the opposite — a marked tendency for people to *underestimate* the odds that a given individual could move up the social ladder. For instance, when we asked our participants ( $N = 865$  MTurkers) to estimate the percentage of children born into the bottom third of the income distribution who failed to ascend at least to the middle class by their mid-20s, their average estimate was 59% (compared to the actual bottom-third immobility rate of 49%; see Chetty et al., 2014). Thus, according to our data (Chambers et al., 2015), there is more upward mobility than most Americans seem to realize (see also Chambers, Swan & Heesacker, 2014).

The authors thank Ryan J. McCarty and Olivia Reyes for their help with manuscript preparation.

Copyright: © 2016. The authors license this article under the terms of the Creative Commons Attribution 3.0 License.

\*Department of Psychology, University of Florida, P.O. Box 112250, Gainesville, FL, USA, 32611. Email: lkswan@ufl.edu.

†Saint Louis University

‡University of Chicago Booth School of Business

The magnitude of the discrepancy between these independent investigations suggests a cause other than the usual replication-failure suspects (e.g., the inherent instability of  $p$ -values within and across studies; see Cumming, 2014) — the two research teams utilized similar measures, employed large samples from comparable participant pools (Davidai and Gilovich reported no meaningful differences between their nationally-representative Harris Poll and MTurk samples), and observed sizeable effects in opposite directions across multiple studies (including exact replications). Moreover, Kraus and Tan (2015) replicated Davidai and Gilovich's overestimation finding independently *and* under conditions of pre-registration (see Kraus, 2015). Why, then, did these studies so dramatically fail to converge on the question of upward mobility perceptions?<sup>1</sup> We set out to investigate.

First, we noted that the two teams had applied different true-mobility-rate *comparators* — Davidai and Gilovich had weighed their participants' guesses against data from the Panel Study of Income Dynamics (<http://psidonline.isr.umich.edu>) as reported by the Pew Economic Mobility Project (2012); whereas we (Chambers et al.) had relied on tax-record data arranged and analyzed by Chetty, Hendren, Kline, Saez and Turner (2014). Indeed, Chetty and his colleagues estimated substantially *more* actual upward mobility in the United States than did Pew, and when we apposed Chetty et al.'s values to Davidai and Gilovich's participant-guess-averages, their overestimation-of-mobility effect disappeared entirely. For instance, participants in Davidai and Gilovich's (2015) Study 1 guessed on average that 33% of people born into the bottom quintile would remain stuck at the bottom later in life — a spot-on estimate of the “true” rate (33%) reported by Chetty et al. Pew estimated the true rate to be 43%. Yet, Davidai and Gilovich's overestimation effect did not reverse — given that our data (Chambers, Swan & Heesacker, 2015) revealed a robust underestimation effect against the same comparator, we reasoned that the use of different accuracy benchmarks alone could not explain the discrepancy between our findings and Davidai and Gilovich's.

Next, assuming that responses elicited from participants are (a) not perfect reflections of their underlying beliefs and (b) often shaped and potentially biased by ostensibly innocuous features of the elicitation method (e.g., Elson, 2016; Slovic, 1995), we turned our attention to differences between the two teams' *stimuli* (Table 1). Despite the fact that the two questionnaire items shared the relatively narrow goal of measuring respondents' perceptions of upward social mobility rates by soliciting numerical estimates, we identified many potentially-meaningful points of departure. For in-

stance, whereas we (Chambers et al., 2015) asked our participants to think about mobility rates *in the past* (how children born to parents in the 1980s are faring *today*), Davidai and Gilovich's (2015) wording might be interpreted as a question about mobility rates *moving forward* (the likelihood that a person born today *will* end up in a given income group *in the future* as an adult). In other words, the two teams may have simply asked and answered different questions about mobility perceptions (if participants indeed perceived this difference), rendering the two results non-comparable. We noted many technical distinctions, too, both in the question *stems* (e.g., the presence/absence of a ladder graphic) and in the formatting of the *response-option sets* (e.g., asking participants to divide the population into “thirds” or into “20%” strata). To bring these myriad differences between the two stimuli under experimental control, we focused first on the latter (the response-option sets), concerned specifically that asking participants to parcel the population into *five* groups (20% segments; Davidai & Gilovich) as opposed to *three* (Chambers et al.; we reasoned that these would be readily identifiable to many Americans as the familiar “upper,” “middle,” and “lower” classes) may have unduly influenced one or both of the two team's outcomes (see Eriksson & Simpson, 2012 for a precedent). Alternatively, if we could rule out this and other response-option-format factors, we would move on to dismantle features of the question stems.

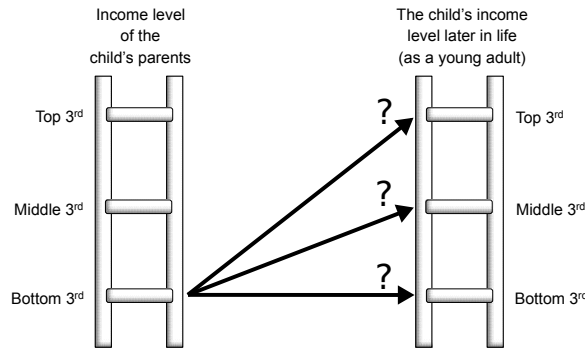
In the following sections, we present the results of new experiments conducted to test this hypothesis that the quintile-tertile factor explains a significant share of the discrepancy between the Davidai-Gilovich (2015) and Chambers et al. (2015) reports. First, we attempted exact replications of key findings reported by both teams. Next, we *inverted* the quintile-tertile factor while holding all other differences in question-stem and response-option wording constant, then assigned a new group of participants at random to one of the two inversion conditions. Finally, we piloted a new question for measuring perceptions of social mobility.

We will *not* conclude this article by declaring one of the two teams' findings the Winner. In our view, the decisions made by both teams of researchers represent the internal-external-validity tradeoffs that all social scientists must make, and, in light of the many differences between the two teams' stimuli (Table 1), it might be reasonable to conclude that the two approaches simply measure different facets of the same social-mobility-perception phenomenon. Moreover, econometric estimates of the *actual* rates of income mobility in the United States often diverge significantly (e.g., see Bloome, 2015, for an overview of current controversies), and *perceptions* of these trends are likely to be moving targets, sensitive to shifting cultural winds. (Note that both teams' studies were completed before primary voting had commenced for the United States' 2016 presidential election.) We conducted the present methodological comparisons with an eye toward refining a *subjective* measure

<sup>1</sup>The two articles did draw consonant conclusions about several other mobility-related topics, including (a) asymmetries between upward and downward mobility estimates, and specifically (b) the tendency for people to underestimate *downward* mobility; and (c) differences in mobility perceptions between conservatives and liberals.

TABLE 1: Full text of the two survey items under investigation.

Chambers, Swan & Heesacker (2015)



Consider a group of American children (born in the early 1980's) to parents in the BOTTOM 3rd of the income distribution, which represents the lowest rung of the "income ladder." In other words, children of "lower class" parents. By the time those children have grown up to be young adults, in their mid-20's, what percentage of them do you think ended up in each of the following income categories? In other words, what percentage of them do you think stayed in the bottom third of the income distribution (i.e., lower class), like their parents, and what percentage of them moved up to the middle third (i.e., middle class) or to the top third (i.e., upper class)?

Please type your estimates for each category in the boxes below. Note: your estimates for the three categories should sum to exactly 100%.

Top third (i.e., upper class)      \_\_\_ %  
 Middle third (i.e., middle class)    \_\_\_ %  
 Bottom third (i.e., lower class)    \_\_\_ %

Davidai and Gilovich (2015)

In the question below, we refer to 5 different economic groups of the population. We divide the U.S. population into five groups, each containing 20% of the total population. These groups are: the richest 20% of the population, the second richest 20% of the population, the middle 20%, the second poorest 20% and the poorest 20% of the population.

We'd like to ask you a question about social mobility in the United States with respect to income. The question below asks you to estimate the chances that the income of an American picked at random would differ from that of his or her parents'. More specifically, when answering these questions, imagine that we took a person born into a family in the poorest 20% of the population at random. What is the likelihood that such a person would be in each of the following income groups as an adult?

[MUST ADD TO 100]

1. Likelihood of a person born into a family in the poorest 20% (as an adult) to the richest 20%      |\_|\_|\_|
2. Likelihood of a person born into a family in the poorest 20% rising (as an adult) to the second richest 20%    |\_|\_|\_|
3. Likelihood of a person born into a family in the poorest 20% rising (as an adult) to the middle 20%      |\_|\_|\_|
4. Likelihood of a person born into a family in the poorest 20% rising (as an adult) to the second poorest 20%    |\_|\_|\_|
5. Likelihood of a person born into a family in the poorest 20% remaining (as an adult) in the poorest 20%    |\_|\_|\_|

of perceived opportunity for socio-economic advancement, which should aid researchers who wish to model its causes and consequences (particularly in reference to some external criterion). Thus, when we refer to *underestimation* or *overestimation* of upward social mobility throughout the rest of this article, we do so solely as a function of our decision to use the Chambers et al. (2015) and Davidai and Gilovich (2015) comparison as a psychometric case study.

## 2 Method

Using Qualtrics © survey hosting software, we constructed a single questionnaire with the following elements: (1) Davidai and Gilovich's (2015) original upward-mobility question prompt; (2) our (Chambers et al., 2015) original upward mobility prompt; (3–4) systematically modified versions of both teams' original items; and (5–6) a new set of items designed to mitigate the quintile-tertile confound that we hypothesized might be driving participants' reactions to the two teams' prompts apart. We assigned participants at random to one of these six survey conditions, which we describe in the Results sections below. The full questionnaire is in the supplement, and also in <https://osf.io/9ya67/> (which includes the .qsl version).

Participant recruitment and data collection occurred on Amazon.com's Mechanical Turk (MTurk) service between March 5th and March 8th, 2015 (before primary voting for the 2016 US presidential election). We paid 2,250 unique MTurk users \$0.50 for completion of the survey — assuming (a) some unknown amount of data loss (e.g., participants who sign up on MTurk but do not actually complete the survey) and (b) that effect sizes tend to shrink in direct replication studies, this large sample size allowed us to detect small effects ( $d = .20$ ) when comparing means between our six experimental conditions with adequate power (.80 with 333 participants per condition). Our MTurk inclusion criteria encompassed (a) identification as a US citizen and (b) a site-wide "HIT" (task) acceptance percentage greater than 95. (The supplement contains the full text of our HIT advertisement.) No participants were excluded from data analysis, which we conducted using IBM's SPSS. (We eliminated missing data in a listwise fashion.)

## 3 Results

**Are Both (Contradictory) Findings Replicable?** First, we performed exact replications of both research teams' upward-mobility-estimate procedures. In the Davidai-Gilovich (DG) condition ( $n = 351$ ), we asked participants to "imagine that we took a person born into a family in the poorest 20% of the population at random," and then to quantify the likelihood that such a person would end up in each of the five income quintile groups in adulthood (see Table

1).<sup>2</sup> Replicating Davidai and Gilovich's original findings, participants in the DG condition estimated on average that only 39.8% ( $Mdn = 40.0$ ,  $SD = 20.18$ ) of Americans remain "stuck" in the bottom quintile — a marginally optimistic appraisal, compared to Davidai and Gilovich's (Pew, 2012) accuracy criterion of 43%;  $t(350) = -2.99$ ,  $p = .003$ ,  $d = .16$  (see the supplement for frequency distributions). Conversely, in the Chambers et al. (CSH) condition ( $n = 367$ ), wherein we asked participants to consider tertiles ("thirds") instead of quintiles ("20%"), the trend reversed: compared to the accuracy criterion of 49% (Chetty et al., 2014), participants on average estimated that many *more* Americans ( $M = 56.4\%$ ,  $Mdn = 60.0$ ,  $SD = 19.74\%$ ) would remain in the bottom third as adults;  $t(366) = 7.17$ ,  $p < .001$ ,  $d = .37$ .

Thus, both effects — that Americans either over- or underestimate social mobility in the U.S., ostensibly depending on (a) the accuracy comparator and on (b) how one asks the question — appear to be genuine, albeit somewhat smaller in magnitude than both original reports suggested. (See Table 2.)

**Replacing Quintiles with Tertiles (and Vice Versa).** In addition to the (1) DG ( $n = 351$ ) and (2) CSH ( $n = 367$ ) conditions, we also allocated participants from our overall sample to (3) a condition in which we modified the DG wording to solicit estimates of upward mobility across *tertiles* rather than *quintiles* (DG-Tertile,  $n = 361$ ); and (4) a condition in which we modified the CSH wording to solicit estimates across *quintiles* rather than *tertiles* (CSH-Quintile,  $n = 363$ ). Holding all other differences between the two questions constant (including the addition of a five-rung-ladder graphic in the CSH-Quintile condition (see supplement), we found that inverting the quintile-tertile factor lead to a powerful reversal: participants in the DG-Tertile condition now *underestimated* mobility [ $M = 57.7\%$ ,  $Mdn = 60.0$ ,  $SD = 20.77$ ; compared to Chetty et al.'s accuracy criterion of 49%,  $t(360) = 8.00$ ,  $p < .001$ ,  $d = .42$ ], whereas participants in the CSH-Quintile condition now judged mobility accurately [ $M = 42.4\%$ ,  $Mdn = 44.0$ ,  $SD = 21.20$ ; compared to the Pew accuracy criterion of 43%,  $t(362) = -0.51$ ,  $p = .61$ ,  $d = .03$ ]. Table S2 in the supplement reveals this pattern in the form of frequency distributions.

Of course, these inverted quintile and tertile response options were still confounded with their respective comparators — thus far, our inferential tests had always compared quintiles to Pew (2012) and tertiles to Chetty et al. (2014). The *real* test of our hypothesis — that participants would respond differently to quintiles than they would to tertiles — required

<sup>2</sup>We elected to focus our analyses on upward *immobility* estimates — the odds of remaining "stuck" in the bottom of the income distribution across the lifespan — in service of simplicity, though we continue to reference Americans' proclivities for overestimating versus underestimating upward *mobility* prospects — the odds of improving one's ranking — in the interest of remaining consistent with the language of both target articles.

TABLE 2: Summary of methodological details and key findings in the present study

Condition	Population segment judged	Instruction text	N	Accuracy comparator	Main findings
DG	Quintiles	DG	351	Pew (2012)	% of individuals remaining in bottom quintile: Mestimate = 39.8% (vs. 43% actual), one-sample $t(350) = -2.99$ , $d = .16$
CSH	Tertiles	CSH	367	Chetty et al. (2014)	% of individuals remaining in bottom tertile: Mestimate = 56.4% (vs. 49% actual), one-sample $t(366) = 7.17$ , $d = .37$
DG-Tertiles	Tertiles	DG	361	Chetty et al. (2014)	% of individuals remaining in bottom tertile: Mestimate = 57.7% (vs. 49% actual), one-sample $t(360) = 8.00$ , $d = .42$
CSH-Quintiles	Quintiles	CSH	363	Pew (2012)	% of individuals remaining in bottom quintile: Mestimate = 42.4% (vs. 43% actual), one-sample $t(362) = -0.51$ , $d = .03$
Quintiles-forced choice	Quintiles	CSH	362	Pew (2012)	% choosing "underestimates" option: 67% % choosing "accurate" option: 23% % choosing "overestimates" option: 11%
Tertiles-forced choice	Tertiles	CSH	363	Chetty et al. (2014)	% choosing "underestimates" option: 75% % choosing "accurate" option: 17% % choosing "overestimates" option: 8%

Note: The prompt soliciting mobility estimates in the DG instruction text conditions read, “We’d like to ask you a question about social mobility in the United States with respect to income. The question below asks you to estimate the chances that the income of an American picked at random would differ from that of his or her parents’. More specifically, when answering these questions, imagine that we took a person born into a family in the poorest 20% [33%] of the population at random. What is the likelihood that such a person would be in each of the following income groups?” In the CSH conditions, the instructions read, “Consider a group of American children (born in the early 1980’s) to parents in the BOTTOM 20% [33%] of the income distribution, which represents the lowest rung of the “income ladder.” In other words, children of “lower class” parents. By the time those children have grown up to be young adults, in their mid-20’s, what percentage of them do you think ended up in each of the following income categories? Participants then entered percentage estimates for each quintile (tertile).

us to standardize the comparator, too. We happen to favor the Chetty et al. (2014) data (note that the Pew Research Center in 2015 updated its estimations methods to bring them more into line with those used by Chetty et al. but our goal in this paper is not to argue for or against a particular estimate of the true rate of upward social mobility in the U.S. We therefore elected to simply approximate mid-points between the Pew and Chetty et al. figures: a 38% immobility rate for quintiles, and a 43% immobility rate for tertiles. To be clear, we leveraged these rough-mid-point estimates as arbitrary benchmarks that would allow us to test an experimental effect (significantly different profiles of responding relative to *any* benchmark criterion between the quintile and tertile response options). In other words, the results of these analyses might reveal whether or not the quintile-tertile effect is, under certain narrow circumstances, strong enough to affect the outcome of a perceptual-accuracy study of this type, but not whether Americans under- or overestimate social mobility rates.

Table 3 presents the results of our critical simulation tests. Participants who divided upward social mobility rates into *tertiles* tended to *underestimate* regardless of the instruction text (CSH versus DG instructions). How participants responded when prompted to think in quintiles, on the other hand, may have depended on the particulars of the question-

stem wording — they tended to *underestimate* in the CSH-Quintile condition, whereas they tended to respond closer to accuracy in the DG-Quintile condition (see Table 3).

**Piloting a New Measure.** Anticipating this quintile-tertile (response-option) effect, we elected to incorporate two additional experimental conditions into our survey with the intention of un-confounding this feature of the DG and CSH response sets. For instance, could we present participants with a task that keeps the goals of the original studies in mind (quantifying mobility estimates) but that changes the basic cognitive process by which participants make their choices (by, say, relieving them of the need to do math)? One solution, we reasoned, would be to change the task from (a) one that involved the active-elicitation of beliefs about income mobility — and the conversion of those beliefs into discrete integers — into (b) one that involved a simple choice between different hypothetical upward-mobility-rate scenarios.

To pilot our idea, we allocated the remainder of our overall sample to one of two *forced-choice* conditions. Within each condition, participants each viewed three separate images of hypothetical “mobility ladders” (see Figure 1) depicting the percentage of individuals born into the bottom quintile (or tertile) who either remained in that quintile (tertile) as adults or who ascended to a higher one. The percentages varied

TABLE 3: Comparing participants' mean estimates of the percentage of Americans who remain "stuck" at the low-end of the socio-economic spectrum across different instruction texts and response options.

Response options	Instruction text		Independent samples t-tests (CSH vs. DG)	Comparisons against arbitrary benchmarks		
	CSH	DG		Test value	CSH	DG
Tertiles	56%	58%	$t(726) = -0.91, p = .37, d = .06$	44%	$t(366) = 12.02, p < .001, d = .63$	$t(360) = 12.56, p < .001, d = .66$
Quintiles	42%	40%	$t(712) = -1.71, p = .09, d = .13$	38%	$t(363) = 3.98, p < .001, d = .21$	$t(350) = 1.65, p = .100, d = .09$

Note. Relative to an arbitrary accuracy criterion (an approximate mid-point-estimate between the upward immobility rates reported by Chetty et al., 2014 and Pew, 2012), participants underestimated mobility in both tertile conditions ( $p$ 's < .001,  $d$ 's = .63–.66). Their (in)accuracy in the quintile conditions depended on the instruction text (we recorded a trend of underestimation using the CSH instructions,  $p < .001, d = .21$ ; and a trend consistent with accuracy using the DG instructions,  $p = .10, d = .09$ ).

across the three images (presented in random order), such that one ladder depicted erroneously low levels of mobility (the “underestimates” option); a second depicted actual levels of mobility in the U.S. (the “accurate” options); and a third depicted erroneously high levels of mobility (the “overestimates” option).

In the Quintiles-Forced-Choice condition ( $n = 360$ ), the overestimates option displayed the mean percentage estimates reported by participants in Davidai and Gilovich’s (2015) Study 1, the accurate option represented their criterion values (Pew Economic Mobility Project, 2012), and the underestimates option represented a mirror image of the overestimates option (calculated by subtracting the value of each quintile’s overestimation-accurate difference from each corresponding “accurate” percentage).<sup>3</sup> The Tertiles-Forced-Choice condition ( $n = 362$ ) followed a similar procedure, except the underestimates option displayed the mean percentage estimates reported by participants in Chambers, Swan, and Heesacker’s (2015) Study 2, the accurate option represented their criterion values (Chetty et al., 2014), and the overestimates option represented a mirror image of the underestimates option (calculated by subtracting the value of each tertile’s underestimate-accurate difference from each corresponding “accurate” percentage).

Figure 1 paints an unequivocal picture: the vast majority of participants (72%) in *both* (quintile and tertile) conditions underestimated social mobility rates. Because we used some elements the CSH question wording — which inadvertently may have primed participants for pessimism — to explain the forced-choice tasks to our participants in both (quintile and tertile) conditions, we subsequently conducted a small ( $N = 99$ ) conceptual replication study of the forced-choice

<sup>3</sup>We adjusted and rounded these “mirror image” values slightly in the Quintile condition so that each rung’s percentage would display a whole number.

procedure with a streamlined prompt: “Please select the image below that represents your best estimate of the actual level of upward social mobility in the United States” (see our supplemental materials for full methodological details). Again, most participants selected the underestimates option in both the Quintile (68.8%) and Tertile (70.6%) conditions.

## 4 Discussion

From a public interest perspective, perceptions of social mobility appear to matter. Leaders within both major U.S. political parties have proffered legislative solutions to the too-little-upward-mobility problem (see, for instance, Barack Obama’s 2015 State of the Union address; and former Florida Governor Jeb Bush’s 2015 “Right to Rise” political action committee), and the rise in *perceived stagnation* of upward mobility (e.g., Pew Research Center, 2012) has been cited as one of the major drivers of Donald Trump’s 2016 electoral victory (e.g., Phillips, 2016). More locally, social scientists have begun to document theoretically important links between mobility perceptions on the one hand and a host of important political judgment and decision-making outcomes on the other. Kraus and Tan (2015), for instance, found in a sample of 751 MTurk workers that younger respondents and respondents with higher perceived socio-economic class both tended to overestimate the number of Americans who manage to improve their social class, suggesting a role for specific varieties of motivated cognition.<sup>4</sup> Similarly, Day

<sup>4</sup>We could just as easily have conducted this entire investigation by comparing our original results (Chambers et al., 2015) with the data collected by Kraus and Tan (2015). Like Davidai and Gilovich, Kraus and Tan (a) described an attempt to quantify Americans’ (in)accurate perceptions of upward social mobility rates; (b) solicited quintile estimates; (c) observed a significant trend of overestimation; (d) published their findings in January of 2015 (in the Journal of Experimental Social Psychology; see also Kraus’ 2015 pre-registered replication report); and (e) used Pew’s (2012) data as their accuracy criterion. We chose Davidai and Gilovich’s paper merely

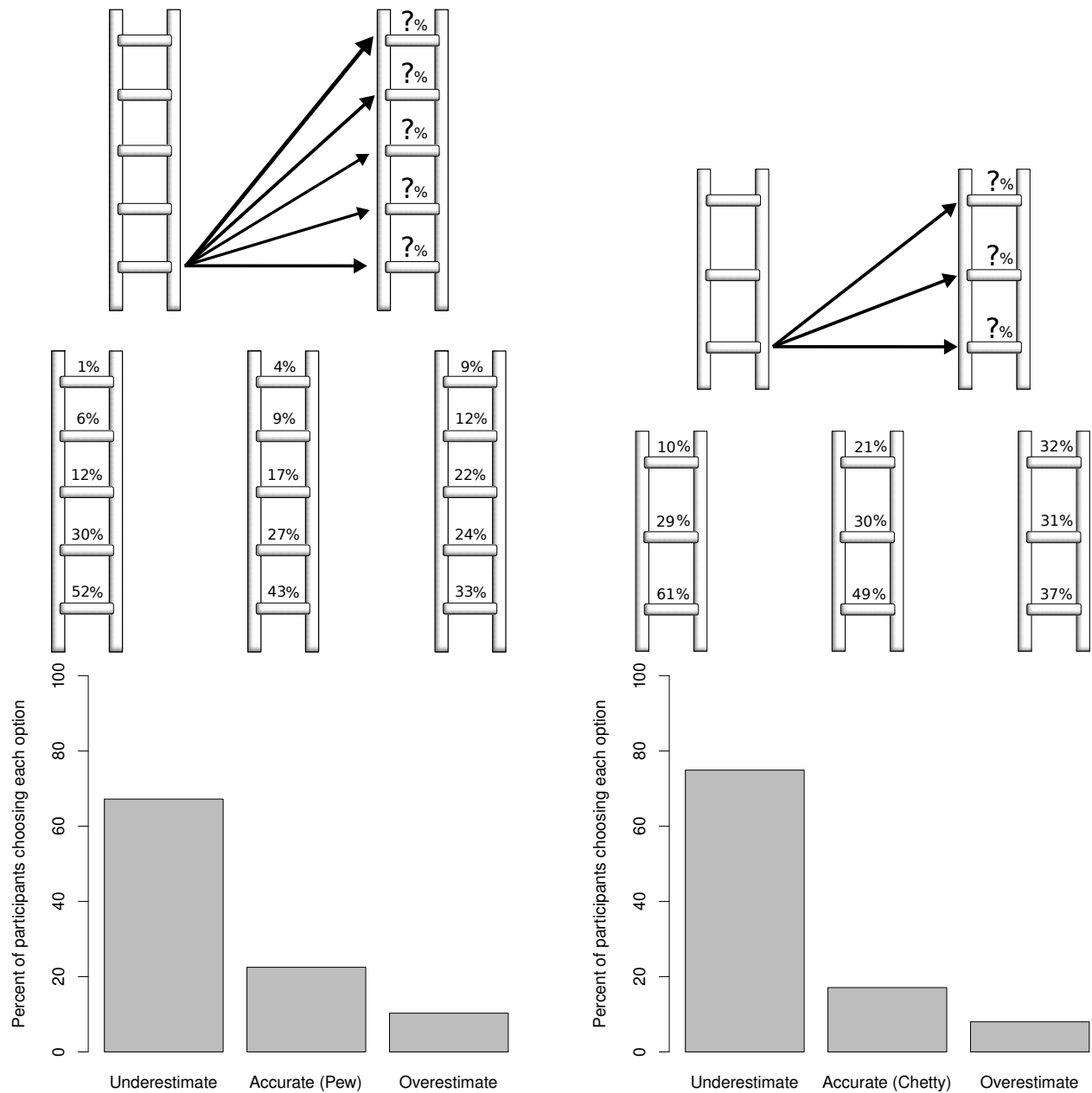


FIGURE 1: Participants (N = 722 MTurk workers), assigned randomly to either a Quintile (five-rungs; left panel) or Tertile (three-rungs; right panel) condition, selected one of three “mobility ladders” to indicate their best estimate of the percentage of individuals born into the bottom (a) quintile (left) or (b) tertile (right) who ended up in each quintile (or tertile) as adults.

and Fiske (2016) found that MTurkers who read a brief report designed to induce perceptions of low upward social mobility subsequently scored lower on the self-report Sys-

tem Justification Scale (Kay & Jost, 2003) relative to an experimental control group (total N = 195). Davidai and Gilovich themselves have produced exciting and provocative data relevant to the questions of (a) *why* some people might in fact overestimate mobility rates relative to their peers (Davidai & Gilovich, 2016a; 2016b); and (b) what sort of upward mobility rates different groups of Americans tend to *prefer* (Davidai & Gilovich, 2015). To be clear, the question-wording measurement error that we identified in the

because we encountered it first. However, we note that Kraus and Tan’s 2015 contribution focused chiefly on comparing mobility perception averages between groups (e.g., between people of high versus low subjective socio-economic status). The question-wording measurement error that we identified in the present experiments need not bear on most of Kraus and Tan’s results, which do not require that the dependent measure be externally valid.

present experiments *need not bear on these findings* — studies that employ participants' social mobility rate estimates as a dependent variable do not require an externally valid measurement in order to obtain reliable and theoretically interesting effects.

We conducted the present investigation with hopes of contributing narrowly to the scholarship surrounding a more specific methodological question: how well do these guess-the-percentage measures capture Americans' knowledge and perceptions of actual social-mobility trends? Following Erikson and Simpson (2012), we wondered about the possibility that the cognitive burden imposed on survey respondents by a quintile/tertile prompt — that is, the task of mentally dividing the American population into fifths or thirds, and then estimating the percentage of people who move between quintiles/tertiles from childhood into adulthood — may have unintentionally biased one or both teams' findings. In other words, we suspected that our original results (Chambers et al., 2015) had diverged so markedly from those reported by Davidai and Gilovich (2015) in part because the two teams had used a different number of response options.

First, to confirm that there really was a reliable discrepancy worth investigating, we replicated both key findings, discovering that, indeed, participants tended to both underestimate (Chambers et al.) and overestimate (Davidai & Gilovich) the true rate of upward social mobility in the U.S., depending on which team's approach (question wording) we used. The most obvious explanation for the divergence, perhaps, was that *because the two results had arisen against different comparators*, they should never have been contrasted in the first place. Yet, when we (a) standardized the comparator and (b) inverted the quintile-tertile factor while holding all other features of the question stems and response sets constant, we still observed significant differences in responding. This effect was sizable enough to alter the binary conclusions that researchers may draw when comparing participant guesses to a known population value, especially when the distance between the accuracy comparator and the average participant's guess is small.

One interpretation of our data is that the quintile approach is more psychometrically fragile than the tertiles approach — it may be the case that when participants cannot access pre-existing attitudes or beliefs about a category (e.g., the “bottom quintile”), they respond more powerfully to other question-stem-related factors, such as (a) differences between population-level “frequencies” (CSH) versus person-specific “likelihoods” (DG); (b) arguably loaded “lower class” labeling in the CSH wording; or, (c) an emphasis on mobility moving forward (DG) rather than mobility in the past (CSH). The preliminary results obtained when we piloted our new measure — a forced-choice survey item that asks participants to select the “true” rate of upward social mobility from a handful of options — support this interpretation. Yet, even if tertile prompts are more resistant to

question-wording artifacts (we would not commit ourselves to this position at this time), participants' responses to those tertile prompts still may not map onto their “true” underlying beliefs about social mobility. In other words, external validity has yet to be established for either approach. We continue to suspect that tertiles more fluidly activate people's associations with “upper,” “middle,” and “lower” classes in America, but this remains speculative.

What can we now say about the (in)accuracy of Americans' perceptions of upward socio-economic mobility rates? Do Americans tend to overestimate them, underestimate them, or perceive them accurately? Our conclusion is that, relying on the available evidence, we simply do not know. Given our observation that researchers' big-picture conclusions in this domain can be swayed by subtle item-wording confounds, one might reasonably conclude that Americans' beliefs about upward social mobility behave more like *constructed preferences* — preferences that are context-dependent and calculated *de novo* at the time of choice (see Warren, McGraw & Van Boven, 2010) — than they behave like fixed attitudes to be extracted. In other words, rather than addressing the question of how Americans think about social mobility, both Davidai and Gilovich (2015) and Chambers et al. (2015) may instead have been addressing more narrow questions regarding how Americans *begin* to think about upward social mobility rates given different initial clues.

What specifically drove the (a) Davidai and Gilovich (2015) and (b) Chambers et al. (2015) papers to such disparate conclusions in the first place? The data we obtained in the present investigation points to a combination of (a) error-inducing bias in the elicitation method (subtle but material differences in question-wording and response formats); and (b) a consequential amount of disagreement about the “true” rate of upward social mobility in the United States (different accuracy comparators). However, if we simply ignored the question of perceptual (in)accuracy, there is virtually no conflict left between the two papers — or between Chambers et al. (2015) on the one hand and Kraus and Tan (2015) and Kraus (2015) on the other — left to resolve.

How should we measure Americans' perceptions of socio-economic mobility moving forward? Contrary to our previous position (Chambers et al., 2015), we now suspect that perceptual (in)accuracy — asking research participants to estimate an estimate that itself is imperfect (economists must infer social mobility rates from incomplete data) — is probably the wrong target for judgment and decision-making scientists in the first place. In light of (a) the complexity involved in designating a gold-standard comparator (Bloome, 2015) and (b) the methodological fragility that we observed in the present experiments, we intend to channel our future efforts away from the question of population-level bias and toward the many important individual-difference questions that remain, including (a) changes in individuals' mobility



perceptions across the lifespan; (b) possible racial, gender, or socio-economic differences (Kraus and Tan, 2015); (c) the differences between laypeople's beliefs about and stated preferences for inequality (another fascinating question raised and addressed by Davidai & Gilovich, 2015); and (d) discrepancies between perceptions of mobility *nationally* versus *locally* (e.g., in one's neighborhood). Each of these phenomena may have separate behavioral implications in different contexts, and each may require a different measurement strategy. When the solicitation of discrete numerical estimates seem appropriate (e.g., when researchers wish to compare participants' perceptions to some external benchmark), we tentatively recommend (a) tertiles over quintiles (quintiles and tertiles clearly produced different patterns of responding in our study, though it may be that neither approach is externally valid); (b) research questions with narrow scopes (e.g., attempts to understand how people's distorted beliefs about how current mobility rates within their community will affect coming generations); and (c) clear theoretical rationales (e.g., a model of the ways in which different varieties of mobility-related beliefs affect voting intentions).

## References

- Bloome, D. (2015). Income inequality and intergenerational income in the United States. *Social Forces*, 93(3), 1047–1080. <https://doi.org/10.1093/sf/sou092>.
- Chambers, J. R., Swan, L. K., & Heesacker, M. (2014). Better off than we know: Distorted perceptions of incomes and income inequality in America. *Psychological Science*, 25(2), 613–618. <http://doi.org/10.1177/0956797613504965>.
- Chambers, J. R., Swan, L. K., & Heesacker, M. (2015). Perceptions of U.S. social mobility are divided (and distorted) along ideological lines. *Psychological Science*, 26(4), 413–423. <http://doi.org/10.1177/0956797614566657>.
- Chetty R., Hendren N., Kline P., Saez E., Turner N. (2014a). *The Equality of Opportunity Project — Online Data Table 1: National 100 by 100 transition matrix*. Retrieved from <http://www.equality-of-opportunity.org/index.php/data>.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. <http://journals.sagepub.com/doi/10.1177/0956797613504966>.
- Davidai, S., & Gilovich, T. (2015). Building a more mobile America — one income quintile at a time. *Perspectives on Psychological Science*, 10(1), 60–71. <http://doi.org/10.1177/1745691614562005>.
- Davidai, S., & Gilovich, T. (2016a). The tide that lifts all focal boats: Asymmetric predictions of ascent and descent in rankings. *Judgment and Decision-Making*, 11(1), 7–20. <http://journal.sjdm.org/15/151021/jdm151021.html>.
- Davidai, S., & Gilovich, T. (2016a). The headwinds/tailwinds asymmetry: An availability bias in assessments of barriers and blessings. *Journal of Personality and Social Psychology*, 111(6), 835–851. <http://dx.doi.org/10.1037/pspa0000066>.
- Day, M. V., & Fiske, S. T. (2016). Movin' on up? How perceptions of social mobility affect our willingness to defend the system. *Social Psychological and Personality Science*. Advance online publication. <http://journals.sagepub.com/doi/abs/10.1177/1948550616678454>.
- Elson, M. (2016). *Question wording and item formation*. In J. Matthes, R. Potter, & C. S. Davis (Eds.), *International Encyclopedia of Communication Research Methods*. Wiley-Blackwell.
- Eriksson, K., & Simpson, B. (2012). What do Americans know about inequality? It depends on how you ask them. *Judgment and Decision Making*, 7(6), 741–745. <http://journal.sjdm.org/12/121027/jdm121027.html>.
- Kay, A. C., & Jost, J. T. (2003). Complementary justice: Effects of “poor but happy” and “poor but honest” stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology*, 85(5), 823–837. <http://doi.org/10.1037/0022-3514.85.5.823>.
- Kraus, M. W. (2015). Americans still overestimate social class mobility: A pre-registered self-replication. *Frontiers in Psychology*, 6, 1709. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4637403>.
- Kraus, M. W., & Tan, J. J. X. (2015). Americans overestimate social class mobility. *Journal of Experimental Social Psychology*, 58, 101–111. <http://doi.org/10.1016/j.jesp.2015.01.005>.
- Obama B. H. (2015, January 20). *State of the Union address*. Retrieved from <https://www.whitehouse.gov/the-press-office/2015/01/20/remarks-president-state-union-address-january-20-2015>.
- Pew Economic Mobility Project. (2012). *Pursuing the American dream: Economic mobility across generations*. Retrieved from [http://www.pewstates.org/uploadedFiles/PCS\\_Assets/2012/Pursuing\\_American\\_Dream.pdf](http://www.pewstates.org/uploadedFiles/PCS_Assets/2012/Pursuing_American_Dream.pdf).
- Pew Research Center. (2012). *The lost decade of the middle class: Fewer, poorer, gloomier*. Retrieved from <http://www.pewsocialtrends.org/2012/08/22/the-lost-decade-of-the-middle-class/>.
- Pew Research Center. (2015). *Economic mobility in the United States*. Retrieved from [http://www.pewtrusts.org/~media/assets/2015/07/fsm-irs-report\\_artfinal.pdf](http://www.pewtrusts.org/~media/assets/2015/07/fsm-irs-report_artfinal.pdf).
- Philips, M. (2016). The hidden economics behind the rise of Donald Trump. *Quartz*. Retrieved from <http://qz.com/626076/the-hidden-economics-behind-the-rise-of-donald-trump/>.
- Right to Rise PAC, Inc. (2015). *The Right to Rise PAC*. Retrieved from <https://righttorisepac.org/>.
- Warren, C. McGraw, A. P., & Van Boven, L. (2011). Values and preferences: Defining preference construction. *Wiley Interdisciplinary Review of Cognitive Science*, 2, 193–205. doi: 10.1002/wcs.98