Journal of Memory and Language 71 (2014) 17–38

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Knowing the crowd within: Metacognitive limits on combining multiple judgments

Scott H. Fraundorf *, Aaron S. Benjamin

University of Illinois at Urbana-Champaign, United States

ARTICLE INFO

Article history:
Received 8 January 2013
revision received 27 August 2013
Available online 13 November 2013

Keywords:
Metacognition
Subjective fluency
Judgment and decision making

ABSTRACT

We investigated how decision-makers use multiple opportunities to judge a quantity. Decision-makers undervalue the benefit of combining their own judgment with an advisor's, but theories disagree about whether this bias would apply to combining several of one's own judgments. Participants estimated percentage answers to general knowledge questions (e.g., *What percent of the world's population uses the Internet?*) on two occasions. In a final decision phase, they selected their first, second, or average estimate to report for each question. We manipulated the cues available for this final decision. Given cues to general theories (the labels *first guess*, *second guess*, *average*), participants mostly averaged, but no more frequently on trials where the average was most accurate. Given item-specific cues (numerical values of the options), metacognitive accuracy was at chance. Given both cues, participants mostly averaged and switched strategies based on whichever yielded the most accurate value on a given trial. These results indicate that underappreciation of averaging estimates does not stem only from social differences between the self and an advisor and that combining general and item-specific cues benefits metacognition.

Introduction

The opportunity to revise a judgment offers both opportunity and challenge. Altering a business projection, reconsidering the accuracy of world knowledge retrieved from memory, or reassessing the time needed to complete a project affords the use of additional information not included in the original judgment. Indeed, making multiple estimates permits greater accuracy in judgment than what could be achieved with a single estimate: the aggregate of multiple estimates, even from the same individual, can outperform any single judgment by reducing the influence of random error on the judgment process (Herzog & Hertwig, 2009; Vul & Pashler, 2008), as detailed below.

However, a judge who has made multiple estimates also faces a decision about how to use those estimates: Is a particular estimate the most accurate; if so, which? Would the estimates be even better if aggregated? Although combining several estimates is generally the most effective strategy (Rauhut & Lorenz, 2010; Vul & Pashler, 2008), the literature suggests that decision-makers often do not make optimal use of multiple estimates. When given the opportunity to choose their own judgment, choose a judgment made by another person, or combine them, judges typically overrely on their own estimates even when judgment accuracy could be improved by combining them (Bonaccio & Dalal, 2006).

Using multiple self-generated estimates does not necessarily present the same challenges as estimates from other judges. One hypothesis is that the bias against combining one's own estimation with others' is due to social factors such as norms on how much advice should be taken or a belief that one is better than the average judge (Harvey & Fischer, 1997). This account does not predict similar underuse of averaging multiple estimates that are all self-generated and do not involve another person. An alternate hypothesis, however, is that suboptimal use of multiple judgments reflects broader cognitive chal-

* Corresponding author. Address: Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL 61820, United States.
E-mail address: scottfraundorf@gmail.com (S.H. Fraundorf).

0749-596X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jml.2013.10.002

Table 1

Across all studies, stimulus questions, correct answers, and mean and standard deviation of participant guesses for all trials; for trials in which the first estimate, second estimate, and average constituted three distinct integer values (75% of all trials); and for trials in which not all response options were distinct integer values (25% of all trials).

	Answer	Estimate 1		Estimate 2		Average	
		M	*SD*	*M*	*SD*	*M*	*SD*
All trials							
Q1: The area of the USA is what percent of the area of the Pacific Ocean?[a]	6.3	28.7	19.6	28.0	19.3	28.4	18.5
Q2: What percent of the world's population lives in either China, India, or the European Union?[a]	44.4	58.0	17.0	61.3	17.7	59.6	15.7
Q3: What percent of the world's airports are in the United States?[a]	30.3	33.4	20.4	34.1	19.7	33.8	18.2
Q4: What percent of the world's roads are in India?[a]	10.5	14.7	14.8	18.3	16.4	16.5	14.2
Q5: What percent of the world's countries have a higher fertility rate than the United States?[a]	58.0	36.4	23.1	37.2	24.6	36.8	21.8
Q6: What percent of the world's telephone lines are in China, the USA, or the European Union?[a]	68.0	72.1	17.9	64.4	21.3	68.2	16.6
Q7: Saudi Arabia consumes what percentage of the oil it produces?[a]	18.9	21.5	19.9	20.0	21.4	20.8	19.9
Q8: What percentage of the world's countries have a higher life expectancy than the United States?[a]	20.3	24.4	19.9	26.1	19.2	25.3	17.9
Q9: What percent of the United States population lives in Florida?	6.0	10.0	7.7	11.6	11.0	10.8	8.7
Q10: What percent of the world's population is 14 years of age or younger?	26.3	32.7	15.2	32.6	17.5	32.6	15.4
Q11: The Internet is used by what percent of the world's population?	30.3	60.6	23.1	58.3	25.0	59.5	23.3
Q12: The European Union consumes what percent of the world's electricity?	16.2	30.2	14.9	33.2	18.4	31.7	14.7
Trials with different first estimate, second estimate, and average (retained for analysis)							
Q1[a]	6.3	29.9	19.1	28.9	18.8	29.4	17.5
Q2[a]	44.4	57.8	17.1	61.8	17.8	59.8	15.4
Q3[a]	30.3	34.0	20.2	35.0	19.3	34.5	17.3
Q4[a]	10.5	16.0	15.6	21.0	17.2	18.5	14.6
Q5[a]	58.0	38.1	23.2	39.0	25.1	38.5	21.7
Q6[a]	68.0	70.8	18.8	61.3	21.9	66.1	16.8
Q7[a]	18.9	23.5	20.9	21.4	23.0	22.5	20.9
Q8[a]	20.3	25.7	20.7	27.9	19.7	26.8	18.2
Q9	6.0	11.2	8.2	13.8	12.4	12.5	9.4
Q10	26.3	33.0	15.2	32.9	18.3	33.0	15.4
Q11	30.3	60.9	22.7	57.8	25.1	59.3	22.9
Q12	16.2	29.8	14.3	33.3	18.5	31.6	14.1
Trials without different first estimate, second estimate, and average (excluded)							
Q1[a]	6.3	26.0	20.6	25.9	20.6	26.0	20.6
Q2[a]	44.4	58.9	17.2	59.0	17.2	59.0	17.2
Q3[a]	30.3	31.4	21.0	31.4	21.0	31.4	21.0
Q4[a]	10.5	11.4	12.0	11.4	12.0	11.4	12.0
Q5[a]	58.0	28.9	20.8	29.0	20.7	29.0	20.8
Q6[a]	68.0	77.7	11.4	77.8	11.5	77.8	11.4
Q7[a]	18.9	17.2	17.0	17.1	17.0	17.1	17.0
Q8[a]	20.3	20.1	16.0	20.1	16.0	20.1	16.0
Q9	6.0	7.7	6.3	7.7	6.2	7.7	6.2
Q10	26.3	31.7	15.4	31.7	15.5	31.7	15.5
Q11	30.3	59.9	24.7	60.0	24.8	59.9	24.7
Q12	16.2	32.4	17.8	32.5	17.8	32.5	17.8

Note: SD = standard deviation.
[a] Item used by Vul and Pashler (2008).

lenges—such as an incorrect belief about the mathematical value of averaging (Soll, 1999) or an overreliance on one's present state of mind—that could impair effective use even of one's own judgments. Thus, investigating how decision-makers use multiple opportunities to estimate the same quantity reveals not only whether and how effectively individuals can apply the normatively correct strategy of combining those estimates, it can also indicate the broader mechanisms by which people make use of multiple, potentially conflicting judgments.

In the present study, we assessed how—and how effectively—decision-makers use several judgments made in response to the same world knowledge question. In particular, we contrast two bases on which participants might decide how to choose or combine those judgments: (a) the plausibility of particular individual estimates and (b) general naïve theories about the value of averaging and of early and later judgments (Soll, 1999). We ask whether metacognition about multiple estimates is more effective given cues supporting one basis or the other—or both together—and what differential performance across cues reveals about the metacognitive bases for such decisions.

The wisdom of crowds and the crowd within

Individuals are frequently called upon to make quantitative estimates, such as projecting a business's sales, forecasting the temperature, judging the time needed to complete a project, or simply answering general knowledge questions such as *What percent of the world's population is 14 years of age or younger?* These estimations have been modeled (Yaniv, 2004) as a function of three sources: (a) the true value, (b) a systematic bias on the part