Julia Minson & Jennifer Mueller

The Wharton School,
University of Pennsylvania

The cost of collaboration:

Why joint decision-making exacerbates rejection of outside information

This paper is part of the Decision Making for a Social World webconference

Organized by the International Cognition and Culture Institute and
the Philosophy, Politics and Economics Program at the University of Pennsylvania

The paper is followed by a **discussion**

Many of the most spectacular institutional catastrophes of the last century are traceable in some way to individually intelligent actors collaboratively making disastrous decisions. Because some of these failures could have been avoided through better use of outside information (Ancona & Caldwell, 1992), researchers have conducted a detailed exploration of group characteristics that give rise to such myopic disregard of alternative viewpoints (Ancona, 1990; Cronin & Weingart, 2007; Janis, 1982; Kane, Argote, & Levine, 2005; Katz, 1982). While research has identified several factors that determine whether collaborators resist or embrace integrating outside information (Ancona, 1990), we pose a novel question. What if, the mere act of collaboration encourages decision makers to resist outside input?

In the current paper, we propose that the greater confidence brought about as a result of making a decision jointly rather than alone, will limit the extent to which collaborators are receptive to outside advice. And while greater confidence is potentially justified by the greater accuracy of collaborative judgments (Minson, Liberman & Ross, 2011), it nevertheless precludes effective integration of outside input and ultimately limits judgment quality.

Research on judgment aggregation has to date exclusively focused on individual level processes (Gino & Moore, 2007; Gino & Schweitzer, 2008; Soll & Mannes, in press). Individuals often improve their decision-making by accepting input from others (see Bonaccio & Dalal, 2006 for a review). This happens in part because they can determine the relative accuracy of own and others' judgments (Soll & Larrick, 2009), and in part due to the statistical truism that aggregating independently-made judgments

reduces average error (Armstrong, 2001; Bonaccio & Dalal, 2006; Einhorn & Hogarth, 1978).

At first blush, one might suggest that collaborative decisions would be more amenable to revision than decisions made by an individual. Individuals may underweight peer input because they are overly attached to their own views (Harvey & Harries, 2004; Lord, Ross, & Lepper, 1979). Collaborative judgments, by definition, require that individuals cede their own views to reach consensus and thus may be less satisfied with, and more open to revising those judgments. Additionally, jointly derived judgments require discussion, which should enable decision makers to selectively integrate superior outside input. By contrast, individuals working alone must rely on their own knowledge to discern the relative quality of their own versus others' judgments.

Alternatively, collaborators may devalue outside input *more* than individuals. Brainstorming research notes that discussion and working collaboratively can increase conformity pressures to maintain the status quo (Goncalo & Staw, 2006). Furthermore, collaborators may choose to disregard outside advice to preserve and even reinforce feelings of cohesion and rapport (Ancona, 1990; Janis, 1982; Katz, 1982).

Additionally, performing a task jointly may increase confidence. Relative to working alone, people believe that working collaboratively allows for acquisition of needed resources, avoidance of negative outcomes, and achievement of desired goals (Moreland, 1987). Correspondingly, research shows that collaboration increases efficacy beliefs including confidence in decision-making (Forsyth, 1999; Park & Hinsz, 2006), and beliefs about overall capability (Stroebe, Diehl, & Abakoumkin, 1992).

In line with the above reasoning, we predict that the tendency to underweight outside input will be more pronounced when decisions are made jointly than when they are made individually. We suggest that relative to individuals, people who make decisions jointly will feel more confident in the accuracy of their responses and thereby more resistant to judgment revision. And although greater confidence may be justified since decisions made collaboratively are likely to be more accurate than those made individually, such confidence may also limit partners' ability to maximize the accuracy of their judgments through sensible use of outside input.

Present Research

In the present research we compare numerical judgments made by individuals with access to the input of another individual to judgments made by dyad members with access to the input of another dyad. We examine the amount of adjustment toward peer input that is made by individuals versus dyad members in light of new information and the effect that this adjustment strategy has on the accuracy of the revised estimates. Additionally, we test whether confidence in estimation accuracy mediates the difference in use of outside input between individuals and dyads.

Method

Participants

Participants (N = 124) were members of a research pool at a large private university who received $10 for participation. Additionally they were offered an accuracy bonus of $30 for each of two estimation rounds, which we reduced by $1 for each percentage point by which any of their estimates deviated from the correct answer.

Procedure

During each session 10-14 participants sat in a large room in front of individual computers separated by partitions and made estimates regarding eight numerical quantities related to U. S. geography, demography, and commerce (Table 1). In the *Individual condition* (n = 36) participants made an initial round of estimates working alone and then revised their estimates after viewing the estimates made by another participant. In the *Dyad condition* (n = 44 dyads) participants worked in pairs to jointly decide on a single set of estimates for each item, and then jointly decided on a set of revised estimates after viewing the estimates produced by another *dyad*.

Table 1: Estimation items and correct answers.

Item wording	Correct Answer
How many accredited medical schools are there in the US?	133
What percentage of the US population are Roman Catholics?	23.9%
What percentage of American households owns dogs?	37.2%
How many people died of HIV/AIDS in the United States in 2008?	22,000
What was the median annual income in 2008 for a US household?	$50,233
How many television broadcast stations are there in the United States?	2218
How long (in miles) is the border between the United States and Mexico?	1969
How many US cities have a population of over 100,000?	273

We calculated the percentage of the distance between the participants' initial estimate and their peers' initial estimate that participants yielded in arriving at their

revised estimates (Bonaccio & Dalal, 2006). Movement in the opposite direction of the peer estimate was coded as 0% and movement beyond the peer estimate was coded as 100%. For each item, participants reported their confidence that their estimate fell within 10 percentage points of the correct answer on a five-point scale anchored at *"Not at all confident"* and *"Extremely confident"*. In the dyad condition we averaged the responses of both dyad members to create a single confidence score.

After the completion of the estimation task all participants provided ratings of cohesion with their dyad partner. The term "dyad partner," meant different things in the two conditions: participants in the dyad condition considered their partner to be the person with whom they worked to make the estimates, whereas the participants in the individual condition considered their partner to be the person whose estimates they viewed. While this difference in construal prevents us from making between-condition cohesion comparisons, the measure was used in within-condition analyses.

Results

Use of peer input

Consistent with prior research on judgment revision, individuals adjusted 35.2% of the distance from their own to the other participant's estimate, giving roughly twice as much weight to their own prior estimates as they did to those of their peers. In line with our predictions, the dyads in our study adjusted even less, yielding an average of 25.3%. In order to examine the significance of this between-condition difference we tested a hierarchical linear model in Stata using individuals as the grouping variable in the individual condition and dyads as the grouping variable in the dyad condition. The between condition difference proved to be significant, $B = .098$, $z = 2.54$, $p < .02$.

Participants in the dyad condition (who actually interacted with their dyad partners) reported higher levels of cohesion (M = 5.99), than did participants in the individual condition who were rating their cohesion with the person whose estimates they received, but with whom they had no other interaction (M = 2.33). However, in neither condition was the level of cohesion predictive of the extent to which participants yielded to peer input.

The mediating role of confidence

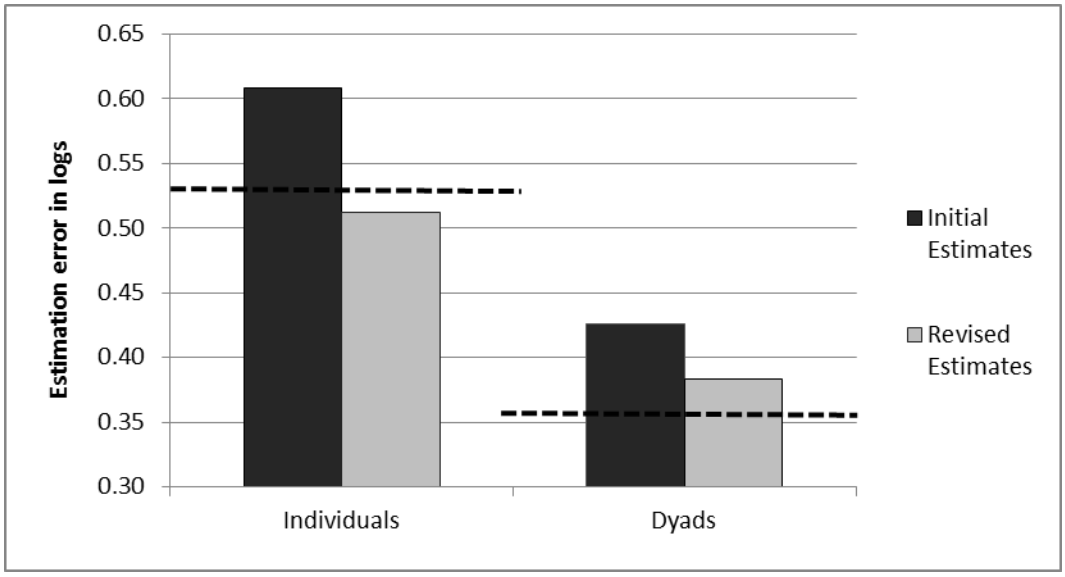
Participants in the dyad condition were more confident in their initial estimates (M = 2.43, SD = 1.02) than participants in the individual condition (M = 1.85, SD = 0.83), B = -.578, z = -4.08, p < 0.001. This pattern persisted (although less dramatically) into the revised estimates (M = 2.69, SD = 0.98 *Dyad condition*; M = 2.24, SD = 0.90 *Individual condition*).

To test whether confidence mediated the relationship between condition and estimate revision we followed the method outlined by Baron and Kenny (1986). Having established that condition predicts use of input, we tested a hierarchical linear model which showed that confidence in initial estimates also predicts our dependent variable, B = -.077, z = -5.24, p < .001. When we regressed yielding to peer input on both experimental condition and confidence in initial estimates, we found that the effect of experimental condition on yielding fell below the level of significance (B = .056, z = 1.41, p = 0.16). The significance of the drop in B for condition from when it was the single predictor in the model to when the mediator was included was confirmed by the Sobel test (B dropped from .095 to .056, z = 3.11, p < .001).

Estimation Accuracy

To examine estimation accuracy, we transformed both the estimates and the correct answer to base 10 logs and then measured estimation error as the absolute difference between the logged estimate and the logged answer. Figure 1 presents the error of initial and revised estimates for both experimental conditions.

Figure 1: Round-by-round estimation error by individuals and dyads. (Dashed lines represent level of accuracy possible by averaging own and peer estimates).



Participants in the dyad condition offered initial estimates that were more accurate (M = 0.426, SD = 0.503) than those in the individual condition (M = 0.608, SD = 0.742), $t(78) = 3.76$, $p < .001$. After receiving input participants in the individual condition achieved an error reduction of M = 0.095, which resulted in revised estimates with an average error of M = 0.513, SD = 0.236. In the dyad condition the average error reduction

of $M = 0.042$ resulted in revised estimates with an error of $M = 0.384$, $SD = 0.160$. Thus, although after revision the estimates of the participants in the dyad condition continued to be more accurate than those in the individual condition, the dyads improved significantly *less* than the individuals ($B = 0.052$, $z = 1.94$, $p = .05$).

To test whether the greater accuracy improvement by individuals versus dyads was a result of the greater use of peer input, we tested a mediation model treating condition as the independent variable, improvement from the first to the second estimate as the dependent variable, and use of peer input as the mediator. Having established that experimental condition predicted accuracy improvement, and that experimental condition predicted yielding to peer input, we tested a hierarchical linear model regressing accuracy improvement on the amount of yielding to peer input and experimental condition.

The results of this analysis showed a significant effect of the extent to which participants yielded to peer input for accuracy improvement ($B = .20$, $z = 4.77$, $p < 0.001$), and a reduced (and not statistically significant) role of experimental condition ($B =.037$, $z = 1.31$, *ns*). The Sobel test yielded a significant statistic (B reduced from .052, to .037, $z = 2.24$, $p < .03$), suggesting that the amount of yielding to peer input indeed fully mediated the difference in accuracy improvement observed between conditions.

Is it possible that the lesser accuracy improvement demonstrated by the dyads was due to a "floor effect," such that having produced more accurate initial estimates, dyad members had less room for improvement? In order to test this alternative explanation we compared the accuracy of the revised estimates in both conditions to the accuracy of estimates that participants *could* have achieved if they had averaged their original estimates with those of their peers (an accuracy benchmark commonly used in the

literature (Minson, Liberman, & Ross, 2010; Soll & Larrick, 2009) and illustrated by the dashed lines in Figure 1).

We tested a hierarchical linear model using the difference between the accuracy of participants' actual revised estimates and the accuracy they could have achieved by averaging as the dependent variable, and experimental condition as the independent variable. We controlled for confidence in initial estimates and error of initial estimates. The results of the model supported our hypothesis. The error of the initial estimates yielded a positive coefficient (B = .16, z = 5.69, $p < 0.001$). Confidence in initial estimates had no significant effect (B = -.03, z = -1.50, *ns*). In line with our predictions, participants in the dyad condition made revised estimates that were significantly less accurate relative to averaging than did participants in the individual condition (B = -.09, z = -2.18, $p < 0.03$). Even controlling for the superior accuracy of initial estimates, participants in the dyad condition underperformed the level of accuracy possible if they had averaged their initial estimates with those of their peers, relative to the participants in the individual condition.

Discussion

Collaboration comes with a price. Specifically, dyads were significantly more reluctant than individuals to revise their judgments and thus paid an accuracy cost. We further found that this unwillingness to yield to outside input was explained by dyad members' greater confidence in their estimates. These findings suggest that discussion and consensus building, the very processes allowing dyads to capitalize on divergent views, may also limit dyads ability to benefit from outside advice.

Our results also point to a paradoxical implication for self-efficacious thinking. Specifically, feelings of confidence or efficacy have been shown to promote performance of individuals and groups alike (Bandura, 1977). Yet, a burgeoning body of research proposes that efficacious feelings do not always promote performance – especially when individuals or groups are engaged in a relatively novel task where feelings of high efficacy may inhibit the types of discussions and exploration which result in improved processes and performance (Goncalo, Polman, & Maslach, 2010; Moore & Healy, 2008). In line with this view, the current investigation shows that efficacious beliefs may also inhibit the extent to which decision-makers consider novel information from outside sources.

A large body of literature shows that knowledge transfer in organizations is difficult because groups are resistant to incorporating novel information (Ancona & Caldwell, 1992a). This work proposes that the quality of collaboration – not the mere act of collaborating explains why members are reluctant to change their mind. The current study suggests that collaborators may resist incorporating outside input more than previously believed as the collaborative process of generating the response increased feelings about the accuracy and appropriateness of that response and accounted for a general resistance to incorporating outside input.

Unlike individuals, dyads were able to discuss estimates, which may have simultaneously increased a feeling of cohesion within the dyad. Prior research suggests that highly cohesive groups may be more self-attentive and disregard outside information (Ancona & Caldwell, 1992b). However, cohesion was not correlated with yielding to outside input. Instead, our findings suggest that rather than wishing to maintain a positive

feeling for the group or conform to group norms, dyads may reject outside information simply because they do not believe the information adds value.

Many of our most important decisions are made collaboratively, following the intuition that "two heads are better than one". Every aspect of law, policy and corporate governance relies on the ability of individuals to jointly make effective decisions. Our study demonstrates that while collaborators do make more accurate judgments, their reluctance to integrate external input can severely impair their ability to achieve their goals.

References

Ancona, D. G. (1990). Outward bound: Strategies for team survival in an organization. *Academy of Management Journal, 33*(2), 334-365.

Ancona, D. G., & Caldwell, D. F. (1992a). Bridging the boundary: External activity and performance in organizational teams. *Administrative Science Quarterly, 37*(4), 634-665.

Ancona, D. G., & Caldwell, D. F. (1992b). Demography and Design: Predictors of New Product Team Performance. *Organization Science, 3*(3), 321-341.

Armstrong, J. (2001). Combining forecasts. In J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners* (pp. 417-440). Norwell, MA: Kluwer Academic Publishers.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes, 101*(2), 127-151.

Cronin, M. A., & Weingart, L. R. (2007). Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review, 32*(3), 761-773.

Einhorn, H., & Hogarth, R. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*(5), 395-416.

Forsyth, D. (1999). *Group dynamics (3rd Ed.).* Belmont, CA: Wadsworth.

Gino, F., & Moore, D. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making, 20*(1), 21-35.

Gino, F., & Schweitzer, M. E. (2008). Blinded by anger or feeling the love: How emotions influence advice taking. *Journal of Applied Psychology, 93*(5), 1165-1173.

Goncalo, J. A., Polman, E., & Maslach, C. (2010). Can confidence come too soon?: Collective efficacy, conflict and group performance over time. *Organizational Behavior and Human Decision Processes, 113*, 13-24.

Goncalo, J. A., & Staw, B. M. (2006). Individualism-collectivism and group creativity. *Organizational Behavior and Human Decision Processes, 100*(1), 96-109.

Harvey, N., & Harries, C. (2004). Effects of judges' forecasting on their later combination of forecasts for the some outcomes. *International Journal of Forecasting, 20*, 391–409.

Janis, I. (1982). *Groupthink*: Houghton Mifflin Boston.

Kane, A. A., Argote, L., & Levine, J. M. (2005). Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality. *Organizational Behavior and Human Decision Processes, 96*(1), 56-71.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative Science Quarterly, 27*(1), 81-104.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098-2109.

Minson, J. A., Liberman, V., & Ross, L. (2010). Two to tango: The effect of collaboration and disagreement on dyadic decision-making. *Working Paper, University of Pennsylvania.*

Moore, D., & Healy, P. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502-517.

Moreland, R. (1987). The formation of small groups. *Review of Personality and Social Psychology, 8*, 80-110.

Park, E. S., & Hinsz, V. B. (2006). "Strength and safety in numbers": A theoretical perspective on group influences on approach and avoidance motivation. *Motivation and Emotion. Special Issue: Approach/Avoidance, 30*(2), 135-142.

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(3), 780-805.

Soll, J. B., & Mannes, A. E. (In press). Judgmental aggregation strategies depend on whether the self is involved. *International Journal of Forecasting*.

Stroebe, W., Diehl, M., & Abakoumkin, G. (1992). The illusion of group effectivity. *Personality and Social Psychology Bulletin, 18*(5), 643-650.

Williams, K. D. (2010). Dyads Can Be Groups (and Often Are) *Small Group Research, 41*, 268-274.