

Originally published 30 September 2010; corrected 1 October 2010. Affiliation note numbers for three authors were corrected.

Scienceexpress

Report

Evidence for a Collective Intelligence Factor in the Performance of Human Groups

Anita Williams Woolley,^{1*} Christopher F. Chabris,^{2,3} Alexander Pentland,^{3,4} Nada Hashmi,^{3,5} Thomas W. Malone^{3,5}

¹Carnegie Mellon University, Tepper School of Business, Pittsburgh, PA 15213, USA. ²Union College, Schenectady, NY 12308, USA. ³Massachusetts Institute of Technology (MIT) Center for Collective Intelligence, Cambridge, MA 02142, USA. ⁴MIT Media Lab, Cambridge, MA 02139, USA. ⁵MIT Sloan School of Management, Cambridge, MA 02142, USA.

*To whom correspondence should be addressed. E-mail: awoolley@cmu.edu

Psychologists have repeatedly shown that a single statistical factor—often called “general intelligence”—emerges from the correlations among people’s performance on a wide variety of cognitive tasks. But no one has systematically examined whether a similar kind of “collective intelligence” exists for groups of people. In two studies with 699 individuals, working in groups of two to five, we find converging evidence of a general collective intelligence factor that explains a group’s performance on a wide variety of tasks. This “c factor” is not strongly correlated with the average or maximum individual intelligence of group members but is correlated with the average social sensitivity of group members, the equality in distribution of conversational turn-taking, and the proportion of females in the group.

As research, management, and many other kinds of tasks are increasingly accomplished by groups—both those working face-to-face and “virtually”(1–3)—it is becoming even more important to understand the determinants of group performance. Over the last century, psychologists made significant progress in defining and systematically measuring intelligence in individuals (4). We have used the statistical approach they developed for individual intelligence to systematically measure the intelligence of groups. Even though social psychologists and others have studied for decades how well groups perform specific tasks (5, 6), they have not attempted to measure group intelligence in the same way individual intelligence is measured—by assessing how well a single group can perform a wide range of different tasks, and using that information to predict how that same group will perform other tasks in the future. The goal of the research reported here was to test the hypothesis that groups, like individuals, do have characteristic levels of “intelligence,” which can be measured and used to predict the groups’ performance on a wide variety of tasks.

Although controversy has surrounded it, the concept of measurable human intelligence is based on a fact that is still as remarkable as it was to Spearman when he first documented it in 1904 (7): People who do well on one mental task tend to do well on most others, despite large variations in the tests’ contents and methods of administration (4). In principle, performance on cognitive tasks could be largely uncorrelated, as one might expect if each relied on a specific set of capacities that was not used by other tasks (8). It could even be negatively correlated, if practicing to improve one task caused neglect of others (9). The empirical fact of general cognitive ability as first demonstrated by Spearman is now, arguably, the most replicated result in all of psychology (4).

Evidence of general intelligence comes from the observation that the average correlation among individuals’ performance scores on a relatively diverse set of cognitive tasks is positive, the first factor extracted in a factor analysis of these scores generally accounts for 30–50% of the variance, and subsequent factors extracted account for substantially less variance. This first factor extracted in an analysis of individual intelligence tests is referred to as “general cognitive ability,” or *g*, and it is the main factor that intelligence tests measure. What makes intelligence tests of substantial practical (not just theoretical) importance is that intelligence can be measured in an hour or less, and is a reliable predictor of a very wide range of important life outcomes over a long span of time, including grades in school, success in many occupations, and even life expectancy (4).

By analogy with individual intelligence, we define a group’s collective intelligence (*c*) as the general ability of the group to perform a wide variety of tasks. Empirically, collective intelligence is the inference one draws when the ability of a group to perform one task is correlated with that group’s ability to perform a wide range of other tasks. Note

that this kind of collective intelligence is a property of the group itself, not just the individuals in it. Unlike previous work that examined the effect on group performance of the average intelligence of individual group members (10), one of our goals is to determine whether the collective intelligence of the group as a whole has predictive power above and beyond what can be explained by knowing the abilities of the individual group members.

The first question we examined was whether collective intelligence—in this sense—even exists. Is there a single factor for groups, a *c* factor, that functions in the same way for groups as general intelligence does for individuals? Or does group performance, instead, have some other correlational structure, such as several equally important but independent factors, as is typically found in research on individual personality (11)?

To answer this question, we randomly assigned individuals to groups and asked them to perform a variety of different tasks (12). In Study 1, 40 three-person groups worked together for up to five hours on a diverse set of simple group tasks plus a more complex criterion task. To guide our task sampling, we drew tasks from all quadrants of the McGrath Task Circumplex (6, 12), a well-established taxonomy of group tasks based on the coordination processes they require. Tasks included solving visual puzzles, brainstorming, making collective moral judgments, and negotiating over limited resources. At the beginning of each session, we measured team members' individual intelligence. And, as a criterion task at the end of each session, each group played checkers against a standardized computer opponent.

The results support the hypothesis that a general collective intelligence factor (*c*) exists in groups. First, the average inter-item correlation for group scores on different tasks is positive ($r = .28$, see Table 1). Next, factor analysis of team scores yielded one factor with an initial eigenvalue accounting for over 43% of the variance (in the middle of the 30–50% range typical in individual intelligence tests), while the next factor accounts for only 18%. Confirmatory factor analysis supported the fit of a single latent factor model with the data ($\chi^2 = 1.66$, $P = .89$ $df = 5$; Comparative-Fit-Index (CFI) = .99, Root-mean-square-error-of-approximation (RMSEA) = .01). Furthermore, when the factor loadings for different tasks on the first general factor are used to calculate a *c* score for each group, this score strongly predicts performance on the criterion task ($r = .52$, $P = .01$). Finally, the average and maximum intelligence scores of individual group members are not significantly correlated with *c* ($r = .19$, *ns*; $r = .27$, *ns*, respectively) and not predictive of criterion task performance ($r = .18$, *ns*; $r = .13$, *ns*, respectively). In a regression using both individual intelligence and *c* to predict performance on the criterion task, *c* has a significant effect ($\beta = .51$, $P = .001$), but average

individual intelligence ($\beta = .08$, *ns*) and maximum individual intelligence ($\beta = .01$, *ns*) do not (see Fig. 1).

In Study 2, we used 152 groups ranging from two to five members. Our goal was to replicate these findings in groups of different sizes, using a broader sample of tasks and an alternative measure of individual intelligence. As expected, this study replicated the findings of Study 1, yielding a first factor explaining 44% of the variance and a second factor explaining only 20%. In addition, a confirmatory factor analysis suggests an excellent fit of the single factor model with the data ($\chi^2 = 5.85$, $P = .32$, $df = 5$; CFI = .98, NFI = .89, RMSEA = .03).

In addition, for a subset of the groups in Study 2, we included five additional tasks, for a total of ten. The results from analyses incorporating all ten tasks were also consistent with the hypothesis that a general *c* factor exists (see Fig. 2). The scree test (13) clearly suggests that a one factor model is the best fit for the data from both studies (Akaike Information Criterion (AIC) = 0.00 for single factor solution). Furthermore, parallel analysis (13) suggests that only factors with an eigenvalue above 1.38 should be retained, and there is only one such factor in each sample. These conclusions are supported by examining the eigenvalues both before and after principal axis extraction, which yields a first factor explaining 31% of the variance in Study 1 and 35% of the variance in Study 2. Multiple-group confirmatory factor analysis suggests that the factor structures of the two studies are invariant ($\chi^2 = 11.34$, $P = .66$, $df = 14$; CFI = .99, RMSEA = .01). Taken together, these results provide strong support for the existence of a single dominant *c* factor underlying group performance.

The criterion task used in Study 2 was an architectural design task modeled after a complex research and development problem (14). We had a sample of 63 individuals complete this task working alone, and under these circumstances, individual intelligence was a significant predictor of performance on the task ($r = .33$; $P = .009$).

But when the same task was done by groups, the average individual intelligence of the group members was not a significant predictor of group performance ($r = .18$, *ns*). When both individual intelligence and *c* are used to predict group performance, *c* is a significant predictor ($\beta = .36$, $P = .0001$), but average group member intelligence ($\beta = .05$, *ns*) and maximum member intelligence ($\beta = .12$, *ns*) are not (see Fig. 1).

If *c* exists, what causes it? Combining the findings of the two studies, the average intelligence of individual group members was moderately correlated with *c* ($r = .15$, $P = .04$), and so was the intelligence of the highest-scoring team member ($r = .19$, $P = .008$). However, for both studies, *c* was still a much better predictor of group performance on the criterion tasks than the average or maximum individual intelligence (see Fig. 1).

We also examined a number of group and individual factors that might be good predictors of *c*. We found that many of the factors one might have expected to predict group performance—such as group cohesion, motivation, and satisfaction—did not.

However, three factors were significantly correlated with *c*. First, there was a significant correlation between *c* and the average social sensitivity of group members, as measured by the "Reading the Mind in the Eyes" test (15); $r = .26, P = .002$. Second, *c* was negatively correlated with the variance in the number of speaking turns by group members, as measured by the sociometric badges worn by a subset of the groups (16); $r = -.41, P = .01$. In other words, groups where a few people dominated the conversation were less collectively intelligent than those with a more equal distribution of conversational turn-taking.

Finally, *c* was positively and significantly correlated with the proportion of females in the group ($r = .23, P = .007$). However, this result appears to be largely mediated by social sensitivity (Sobel $z = 1.93, P = .03$), since (consistent with previous research) women in our sample scored better on the social sensitivity measure than men; $t(441) = 3.42, P = .001$. In a regression analysis with the groups for which all three variables (social sensitivity, speaking turn variance, and percent female) were available, all had similar predictive power for *c*, though only social sensitivity reached statistical significance ($\beta = .33, P = .05$) (12).

In summary, these results provide substantial evidence for the existence of *c* in groups, analogous to a well-known similar ability in individuals. Importantly, this collective intelligence factor appears to depend upon both the composition of the group (e.g., average member intelligence) and also on factors that emerge from the way group members interact when they are assembled (e.g., their conversational turn-taking behavior) (17, 18).

These findings raise many additional questions. For example, could a short collective intelligence test predict a sales team's or a top management team's long-term effectiveness? More importantly, it would seem to be much easier to raise the intelligence of a group than an individual. Could a group's collective intelligence be increased by, for example, better electronic collaboration tools?

Many previous studies have addressed questions like these for specific tasks. But by measuring the effects of specific interventions on a group's *c*, one can predict the effects of those interventions on a wide range of tasks. Thus the ability to measure collective intelligence as a stable property of groups provides both a significant economy of effort and a range of new questions to explore in building a science of collective performance.

References and Notes

1. S. Wuchty, B. F. Jones, B. Uzzi, *Science* **316**, 1036 (2007).
2. T. Gowers, M. Nielsen, *Nature* **461**, 879 (2009).
3. J. R. Hackman, *Leading teams: Setting the Stage for Great Performances* (Harvard Business School Press, Boston, 2002).
4. I. J. Deary, *Looking Down on Human Intelligence: From Psychometrics to the Brain* (Oxford Univ. Press, New York, 2000).
5. J. R. Hackman, C. G. Morris, in *Small Groups and Social Interaction*, Volume 1. H. H. Blumberg, A. P. Hare, K. V. M. Davies, Eds. (Wiley, Chichester, UK, 1983), pp. 331–345.
6. J. E. McGrath, *Groups: Interaction and Performance* (Prentice-Hall, Englewood Cliffs, NJ, 1984).
7. C. Spearman, *American Journal of Psychology* **15**, 201 (1904).
8. C. F. Chabris, in *Integrating the Mind: Domain General Versus Domain Specific Processes in Higher Cognition*. M. J. Roberts, Ed. (Psychology Press, Hove, UK, 2007), pp. 449–491.
9. C. Brand, *The g Factor* (Wiley, Chichester, UK, 1996).
10. D. J. Devine, J. L. Philips, *Small Group Research* **32**, 507 (2001).
11. R. R. McCrae, P. T. Costa, *Journal of Personality and Social Psychology* **52**, 81 (1987).
12. Materials and methods are available as supporting material on Science Online.
13. R. B. Cattell, *Multivariate Behavioral Research* **1**, 245 (1966).
14. A. W. Woolley, *Organization Science* **20**, 500 (2009).
15. S. Baron-Cohen, S. Wheelwright, J. Hill, Y. Raste, I. Plumb, *The Journal of Child Psychology and Psychiatry and Allied Disciplines* **42**, 241 (2001).
16. A. Pentland, *Honest Signals: How They Shape Our World* (Bradford Books, Cambridge, MA, 2008).
17. L. K. Michaelsen, W. E. Watson, R. H. Black, *Journal of Applied Psychology* **74**, 834 (1989).
18. R. S. Tindale, J. R. Larson, *Journal of Applied Psychology* **77**, 102 (1992).
19. This work was made possible by financial support from the National Science Foundation (grant IIS-0963451), the Army Research Office (grant #56692-MA), the Berkman Faculty Development Fund at Carnegie Mellon University, and from Cisco Systems, Inc. through their sponsorship of the MIT Center for Collective Intelligence. We would especially like to thank Stephen Kosslyn for his invaluable help in the initial conceptualization and early stages of this work and Ishani Aggarwal and Wen Dong for substantial help with data collection and analysis. We are also grateful for comments and research assistance from Linda Argote, Emily Anderson, Jonathan Chapman, Margaret Ding,

Snehal Gaikwad, Catherine Huang, Josh Introne, Chris Lee, Natasha Nath, Stuti Pandey, Nate Peterson, Helen Ra, Cary Ritter, Fei Sun, Emily Sievers, Kenzan Tenabe, and Roger Wong. The hardware and software used in collecting sociometric data are the subject of an MIT patent application and will be provided for academic research via a not-for-profit arrangement through Prof. Pentland. In addition to the affiliations listed above, Thomas Malone is also a member of the Strategic Advisory Board at InnoCentive, Inc.; a director at Seriosity, Inc., and chairman at Phios Corporation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1193147/DC

Materials and Methods

Tables S1 to S4

References

2 June 2010; accepted 10 September 2010

Published online 30 September 2010;

10.1126/science.1193147

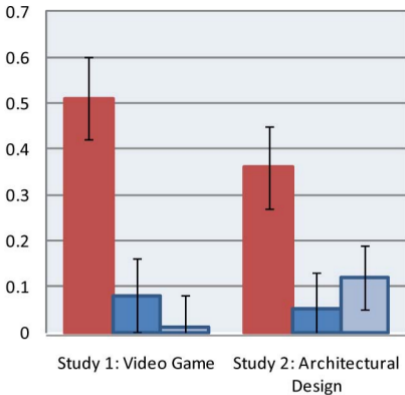
Include this information when citing this paper.

Fig.1. Standardized regression coefficients for collective intelligence (c) and average individual member intelligence when both are regressed together on criterion task performance in Studies 1 and 2 (controlling for group size in Study 2). Coefficient for maximum member intelligence is also shown for comparison, calculated in a separate regression because it is too highly correlated with individual member intelligence to incorporate both in a single analysis; $r = .73$ and $.62$ in Studies 1 and 2.)

Fig. 2. Scree plot demonstrating the first factor from each study accounting for over twice as much variance as subsequent factors. Factor analysis of items from the Wonderlic Personnel Individual Intelligence Test administered to 642 individuals is included as a comparison.

Table 1. Correlations among groups tasks and descriptive statistics for Study 1. $n = 40$ groups; $*P \leq .05$; $**P \leq .001$.

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Collective Intelligence (<i>c</i>)									
2. Brainstorming	.38*								
3. Group Matrix Reasoning	.86**	.30*							
4. Group Moral Reasoning	.42*	.12	.27						
5. Plan Shopping Trip	.66**	.21	.38*	.18					
6. Group Typing	.80**	.13	.50**	.25*	.43*				
7. Avg Member intelligence	.19	.11	.19	.12	-.06	.22			
8. Max Member intelligence	.27	.09	.33*	.05	-.04	.28	.73**		
9. Video Game	.52*	.17	.38*	.37*	.39*	.44*	.18	.13	
Minimum	-2.67	9	2	32	-10.80	148	4.00	8.00	26
Maximum	1.56	55	17	81	82.40	1169	12.67	15.67	96
Mean	0	28.33	11.05	57.35	46.92	596.13	8.92	11.67	61.80
SD	1.00	11.36	3.02	10.96	19.64	263.74	1.82	1.69	17.56

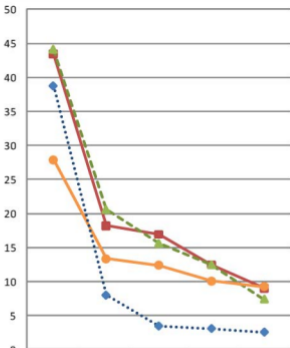


■ Collective Intelligence

■ Average Member Intelligence

■ Maximum Member Intelligence

Percent Variance Explained



	1	2	3	4	5
Study 1	43.39	18.18	16.93	12.46	9.04
Study 2 (five tasks)	44.07	20.54	15.57	12.44	7.39
Study 2 (ten tasks)	27.86	13.35	12.40	10.08	9.24
Indiv Intelligence Test	38.77	8.01	3.47	3.11	2.58