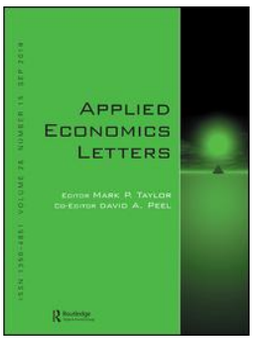

Routledge
Taylor & Francis Group

Applied Economics Letters

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rael20

Teams do not outperform individuals in a simple creative task

Uri Gneezy, Katharina Laske & Marina Schröder

To cite this article: Uri Gneezy, Katharina Laske & Marina Schröder (2021): Teams do not outperform individuals in a simple creative task, Applied Economics Letters, DOI: 10.1080/13504851.2021.1934386

To link to this article: https://doi.org/10.1080/13504851.2021.1934386

APPLIED ECONOMICS LETTERS
https://doi.org/10.1080/13504851.2021.1934386

Routledge
Taylor & Francis Group

Teams do not outperform individuals in a simple creative task

Uri Gneezy[a], Katharina Laske[b] and Marina Schröder[c]

[a]Rady School of Management, University of California San Diego; [b]Faculty of Management, Economics and Social Sciences, University of Cologne; [c]Institute of Economic Policy, Leibniz University Hannover

ABSTRACT

We compare the performance of teams with that of individuals in a simple creative task – generating a title for a short video. To measure performance, we assess the quality of titles using click rate as well as subjective assessment of the fit between the title and the video. Although teams are costlier to organizations, we find no significant differences in the performance of teams relative to that of individuals. As a result, in the task we use in this paper, allocating creative work to individuals is more efficient than allocating it to teams.

KEYWORDS

Creativity; idea generation; real effort; experiment; teams vs. individuals

JEL CLASSIFICATION

C90; O30; J24; M50

1. Introduction

Researchers often argue co-location is important for creative performance, for example because teamwork tends to decrease when the physical distance between team members increases (Powell and Giannella 2010; Paulus, Dzindolet, and Kohn 2012). Recently, COVID-19 forced people all over the world to maintain physical distance from their colleagues, leading to a reduction in time spent communicating in teams (DeFilippis et al. 2020). We provide experimental evidence comparing the performance of teams with that of individuals working on a simple creative task.

Our creative task consists of generating a title for a short video.[1] In a between-subjects design, we compare the quality of titles generated by teams of three with the quality generated by individuals. The strength of the design is that we assess the quality of creative ideas in a way that is natural to the task: we measure the quality of titles through a real click rate (objective measure) and through a subjective evaluation of the fit of the title to the video.

We find that individuals are more likely than teams to provide valid titles (i.e. titles that are related to the content of the video). The average quality of valid titles generated by teams is not significantly different from the quality of valid titles generated by individuals. Further, individuals are not less likely than teams to generate excellent titles. These findings are robust even when we vary the incentives provided for cooperation in teams. Thus, we find no evidence that the creative performance of individuals is inferior to that of teams. When taking costs into consideration, we find that – in our context – individuals are more efficient than teams.

Our research is related to research on brainstorming, which shows teams have fewer and, on average, lower-quality ideas than do individuals (Mullen, Johnson, and Salas 1991). Reasons for the lower performance of teams are productivity blocking, evaluation apprehension, and a lack of individual incentives (Mullen, Johnson, and Salas 1991). Regarding selection of good ideas, however, teams are shown to outperform individuals (Girotra, Terwiesch, and Ulrich 2010; Singh and Fleming 2010).

Our paper also contributes to the recent literature in experimental economics on incentives and idea generation (Eckartz, Kirchkamp, and Schunk 2012; Ederer and Manso 2013; Erat and Gneezy 2016; Bradler, Neckermann, and Warnke 2019; Charness and Grieco, 2020), the dimensions that should be incentivized (Kachelmeier, Reichert, and Williamson 2008; Laske and Schröder 2018), expertise and idea generation (Gneezy, Laske, and

CONTACT Marina Schröder schroeder@wipol.uni-hannover.de

[1]These types of tasks are a big business; for instance, the world's 10 biggest advertising agencies had a combined gross income of more than $25 billion in 2017, with many of their services revolving around creativity (https://co.agencyspotter.com/50-largest-marketing-companies-in-the-world).

Schröder 2021), and how incentives interact with the type of ideation task (Charness and Grieco 2018).

2. Experimental design

The experimental task mimics a creative task in marketing. We asked 543 participants to come up with creative titles for a short video,[2] where the titles should be related to the video. We told participants that the quality of titles would be evaluated according to their title's click-rate success with other participants (see Appendix B for instructions). We classified titles that were completely unrelated to the video as invalid. We programmed the laboratory experiment using ZTree (Fischbacher 2007).

Between treatments, we varied whether participants worked on the task individually or in teams of three (communicating by chat). Because team incentives could affect creativity (Englmaier et al. 2018, 2021; Morgan, Neckermann, and Sisak 2020), we also varied the incentives in teams. In all treatments, participants received a fixed show-up fee of 5. USD In the *Individuals* treatment, participants worked individually, whereas they worked in teams in all other treatments. In the *Individuals* and the *Teams* treatments, participants received a fixed amount of 4 USD for performing the task. In the *Equal Share* and the *Contributor Bonus* treatments, two randomly matched teams competed with each other, and the team that provided the title with a higher click rate received a 24 USD prize in addition to the show-up fee. In the *Equal Share* treatment, the bonus was equally divided among the three team members, so each member of the winning team received 8. USD In the *Contributor Bonus* treatment, the team member who entered the final title into the system received a higher fraction of the prize ($14), whereas the other two team members received only 5 USD each. Importantly, the expected payoff was always 4 USD ($9 including the show-up fee) and did not vary between treatments. We collected data in two waves. In the first wave, conducted in October and November 2016, we aimed at collecting 30 independent observations per treatment. In the second wave, conducted in October 2017 until February 2018, we collected additional data for the teams and individuals treatments. In the main paper, we report results for the full sample. In the Appendix, we additionally report results only for data collected in the first wave. All sessions were run at the Rady Behavioural Lab at the University of California, San Diego. Student participants were recruited through the Lab's subject pool. Each session lasted approximately 30 minutes and the level of payment is in line with the lab standards.

To measure the quality of titles, we recruited 1240 MTurk raters who were blind to the treatment. We provided each rater with four randomly drawn titles (drawn from all titles generated throughout the entire experiment) and asked them to click on one title corresponding to the video they wanted to watch. Raters knew they would subsequently watch the corresponding video and answer some questions about the content of the video. Thus, raters in our experiment made real decisions involving the opportunity cost of time. On average, 17.5 raters saw each title, each time in a different combination. We derive the quality of each title as the fraction of raters who clicked on it; thus, quality ranges from 0 to 1. After raters watched the video, we asked them to assess the fit of the title for the video on a 5-point scale (1 = *very poor* to 5 = *very good*). Average ratings constitute our measure of the fit. Note five titles were never chosen, and thus are excluded in this measure. The rating of titles was performed using the software SoSciSurvey (Leiner 2014).

3. Results

Table 1 provides summary statistics on the fraction of valid titles generated, the click rate of titles, and the fit of titles to the video for the four treatments separately.[3] We find the fraction of invalid titles is larger in the *Teams* treatment than in the *Individuals* treatment (Fisher's exact test, $p < .01$).[4] We find a similar effect when comparing the *Equal Share* and the *Individuals* treatments (Fisher's exact

[2]See https://www.youtube.com/watch?v=GwgtwY3oL4g&feature=youtu.be for the video.
[3]In Table A1 in the appendix, we provide an analysis based only on the first wave of data collection.
[4]We always report two-tailed p-values.

Table 1. Average Scores.

	N (n)	Fraction of Valid Titles (180)	Avg. Click Rate	Avg. Fit
Teams	60		0.85 (0.36) ***	0.24 (0.12) n.s.
3.78 (0.80) n.s.				
Equal Share	27 (81)	0.85 (0.36)**	0.23 (0.14) n.s.	3.71 (0.73) n.s.
Contributor Bonus	30 (90)	0.90 (0.31) n.s.	0.24 (0.12) n.s.	3.71 (0.90) n.s.
Individuals	191	0.96 (0.19)	0.26 (0.12)	3.77 (0.75)

Mean values with standard deviation in parentheses.
N refers to the number of independent observations, and n to the number of subjects.
Invalid titles are excluded in the click rate and fit. Results from Fisher's exact test for the fraction of valid titles, and for U-tests for the click rate and the fit comparing the Individuals treatment with the three Teams treatments are reported in superscript where *** p < .01; ** p < .05; n.s. p > .10.

Table 2. Excellent Titles (Top Quartile).

	Excellent Click Rate	Excellent Fit	Excellent Product
Teams	0.22 n.s	0.25 n.s	0.22 n.s
Equal Share	0.26 n.s	0.07**	0.23 n.s
Contributor Bonus	0.20 n.s	0.37 n.s	0.17 n.s
Individuals	0.26	0.30	0.28

Excellence refers to a dummy equal to 1 if a title is in the top quartile with respect to the click rate (excellent click), the fit (excellent fit), or the product of the click rate and fit (excellent product). Note that quartiles are determined including non-valid illustrations. Results from Fisher's exact tests comparing the Individuals treatment with the three Teams treatments are reported in superscript where ** p < .05; n.s. p > .10. We also conduct binomial tests and observe significant differences to 0.25 only for fit in the Equal Share (Binomial test, p = .04).

test, p = .03), but find no significant difference between the *Contributor Bonus* and the *Individuals* treatment (Fisher's exact test, p = .14).

We find the average click rate of valid titles is slightly but insignificantly higher in the *Individuals* than in the *Teams* treatment (U-test, p = .47). Results are similar when we compare the *Equal Share* or the *Contributor Bonus* treatments with the *Individuals* treatment (U-test, p > .32). We find no significant treatment differences in the average fit of titles generated (pairwise U-test, p > .69). We also find no significant treatment differences in the variance of the click rate or the fit (pairwise variance-ratio test, p > .18).

In our creative tasks, individuals did not perform less well than teams. However, average performance may not be the right measure when assessing creative performance. Firms in this creative context may be seeking a high number of excellent ideas (see, e.g. Girotra, Terwiesch, and Ulrich 2010). To consider this fact, we also analyse positive outliers – 'excellent ideas.' We classify a title as excellent whenever it belongs to the best quartile with respect to the click rate, the fit, or the product of click rate and fit. (See Appendix A for an alternative approach of considering ideas within the 90th percentile.) Table 2 provides an overview of the fraction of excellent ideas for each treatment separately and for the three different measures of excellence. Again, we find no evidence that individuals perform less well than teams. In fact, we find teams in the *Equal Share* treatment are less likely than individuals to produce excellent ideas with respect to the fit (Fisher's exact test, p = 0.01).

The cost of generating a title should also be considered. The human-resource costs of generating titles through individuals are lower than the costs of having teams generate titles. Figure 1

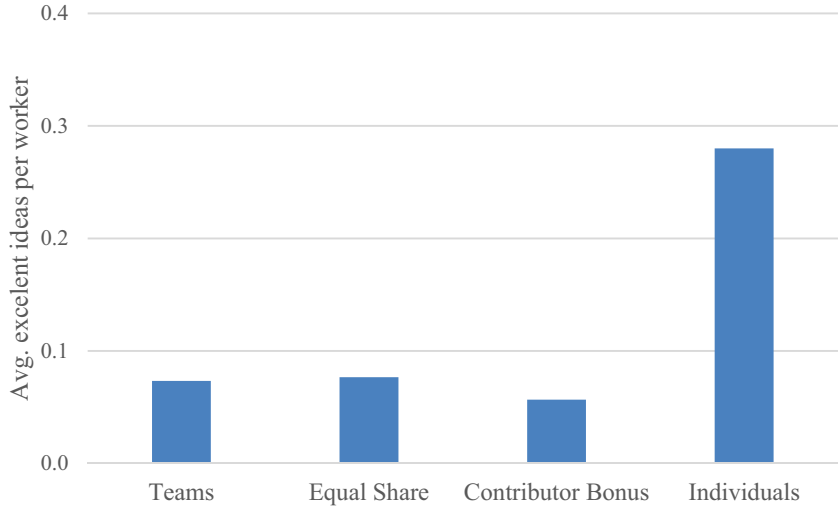

Figure 1. Excellent Ideas per Worker.

depicts the average number of excellent ideas per worker (with respect to the product of fit and click rate) in each of the treatments. In this measure, individual work clearly outperforms teamwork. We find a similar pattern for excellence in the click rate and in fit separately (see Figures A1 and A2 in Appendix A).

4. Conclusion

We study performance in a simple creative task of generating titles for a short video. We find teams do not outperform individuals and that this finding is robust to different incentive schemes used to incentivize teams. Thus, in our simple creative tasks, organizations can increase efficiency by assigning such tasks to individuals instead of teams. It is left to future research to analyse whether this is also the case for more complex creative tasks.

Disclosure of potential conflicts of interest

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the German Research Foundation [FOR 1371].

ORCID

Marina Schröder http://orcid.org/0000-0002-3265-6317

References

Bradler, C., S. Neckermann, and A. J. Warnke. 2019. "Incentivizing Creativity: A Large-Scale Experiment with Performance Bonuses and Gifts." *Journal of Labor Economics* 37 (3): 793–851. doi:10.1086/702649.

Charness, G., and D. Grieco. 2018. "Creativity and Incentives." *Journal of the European Economic Association* 17 (2): 454–496. doi:10.1093/jeea/jvx055.

Charness, G., and D. Grieco. 2020. "Creativity and Corporate Culture." Available at SSRN 3723865.

DeFilippis, E., S. M. Impink, M. Singell, J. T. Polzer, and R. Sadun. 2020. "Collaborating during Coronavirus: The Impact of COVID-19 on the Nature of Work". National Bureau of Economic Research No. W27612.

Eckartz, K., O. Kirchkamp, and D. Schunk. 2012. "How Do Incentives Affect Creativity?" Available at SSRN 2198760.

Ederer, F., and G. Manso. 2013. "*Is Pay for Performance Detrimental to Innovation?*" *Management Science* 59: 1496–1513.

Englmaier, F., S. Grimm, D. Grothe, D. Schindler, and S. Schudy. 2021. "The Efficacy of Tournaments for Non-Routine Team Tasks." Mimeo.

Englmaier, F., S. Grimm, D. Schindler, and S. Schudy. 2018. *"The Effect of Incentives in Non-Routine Analytical Teams Tasks-Evidence from a Field Experiment."*

Erat, S., and U. Gneezy. 2016. "Incentives for Creativity." *Experimental Economics* 19 (2): 269–280. doi:10.1007/s10683-015-9440-5.

Fischbacher, U. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10 (2): 171–178. doi:10.1007/s10683-006-9159-4.

Girotra, K., C. Terwiesch, and K. T. Ulrich. 2010. "Idea Generation and the Quality of the Best Idea." *Management Science* 56 (4): 591–605. doi:10.1287/mnsc.1090.1144.

Gneezy, U., K. Laske, and M. Schröder. 2021. "Creative Solutions: Expertise versus Crowd Sourcing." Mimeo.

Kachelmeier, S. J., B. E. Reichert, and M. G. Williamson. 2008. "Measuring and Motivating Quantity, Creativity, or Both." *Journal of Accounting Research* 46 (2): 341–374. doi:10.1111/j.1475-679X.2008.00277.x.

Laske, K., and M. Schröder. 2018. "Quality through Quantity - the Effects of Piece-Rate Incentives on Creative Performance." CGS Working Paper 7(1).

Leiner, D. J. 2014. "SoSci Survey (Version 2. 5.00-i)[Computer Software]."

Morgan, J., S. Neckermann, and D. Sisak. 2020. "Peer Evaluation and Team Performance: An Experiment on Complex Problem Solving." Mimeo.

Mullen, B., C. Johnson, and E. Salas. 1991. "Productivity Loss in Brainstorming Groups: A Meta-analytic Integration." *Basic and Applied Social Psychology* 12 (1): 3–23. doi:10.1207/s15324834basp1201_1.

Paulus, P. B., M. Dzindolet, and N. W. Kohn. 2012. "Collaborative creativity—Group Creativity and Team Innovation." In *Handbook of Organizational Creativity*, Edited by M. D. Mumford, 327–357. Academic Press. doi.

org/10.1016/B978-0-12-374714-3.00014-8 .

Table A1. Average Scores Only First Wave of Data collection.

	N (n)	Fraction of Valid Titles	Avg. Click Rate	Avg. Fit
Teams	24	(72)	0.75 (0.44) $^{n.s.}$	0.28 (0.14) $^{n.s.}$ 3.40 (0.75) $^{n.s.}$
Equal Share	27	(81)	0.85 (0.36)$^{n.s.}$	0.23 (0.14) $^{n.s.}$ 3.71 (0.73) $^{n.s.}$
Contributor Bonus	30	(90)	0.90 (0.31) $^{n.s.}$	0.24 (0.12) $^{n.s.}$ 3.71 (0.90) $^{n.s.}$
Individuals	41	0.90 (0.30)	0.27 (0.12)	3.70 (0.89)

Mean values with standard deviation in parentheses.
N refers to the number of independent observations, and n to the number of subjects.
Invalid titles are excluded in the click rate and fit. Results from Fisher's exact test for the fraction of valid titles, and for U-tests for the click rate and the fit comparing the Individuals treatment with the three Teams treatments are reported in superscript where *** p < .01; ** p < .05; n.s. p > .10.

Powell, W. W., and E. Giannella. 2010. "Collective Invention and Inventor Networks." In *Handbook of the Economics of Innovation*. Edited by B. H. Hall and N. Rosenberg, Vol. 1, 575–605. North-Holland, Oxford.

Singh, J., and L. Fleming. 2010. "Lone Inventors as Sources of Breakthroughs: Myth or Reality?" *Management Science* 56 (1): 41–56. doi:10.1287/mnsc.1090.1072.

Table A2. Excellent Titles (Top 10%).

	Excellent click	Excellent fit	Excellent product
Teams	0.12 $^{n.s}$	0.12 $^{n.s}$	0.07 $^{n.s}$
Equal Share	0.15 $^{n.s}$	0.04 $^{n.s}$	0.04 $^{n.s}$
Contributor Bonus	0.07 $^{n.s}$	0.13 $^{n.s}$	0.14 $^{n.s}$
Individuals	0.10	0.10	0.12

Excellence refers to a dummy equal to 1 if a title is among the top 10% with respect to the click rate (excellent click), the fit (excellent fit), or the product of click rate and fit (excellent product). Note that percentiles are determined including non-valid illustrations.
Results from Fisher's exact tests comparing the Individuals treatment with the three Teams treatments are reported in superscript where n.s. p > .10. We also conduct binomial tests and observe no significant differences to 0.10 (Binomial test, p > .34).

Appendix A: Additional Analysis

Appendix B: Instructions

Welcome to today's experiment. Please read the instructions carefully. If you have a question, please raise your hand and one of us will come to your desk to answer it. Each participant will be paid a $5 show-up fee.

–next screen –

You are **Participant X.**

–next screen–

Your Task

[Individual Treatment]

In the following you will watch a short video. Your task is to come up with a title for this video. The aim is to find a creative title that generates a high click rate and is related to the video.

You will be randomly matched with another participant from this experiment. We will post the title to people on the internet, and observe which title generates more clicks by people who need to watch one of them. The one that will generate the title with the higher click-rate will win.

You will have 15 minutes to come up with a title. If you do not submit a title before the time is over, you will not receive payment for performing this task.

[Group Treatments]

In the following, you will be assigned to work with two other participants in a group of three. Each one of you will watch the same short video. Your group task is to come up with a title for this video.

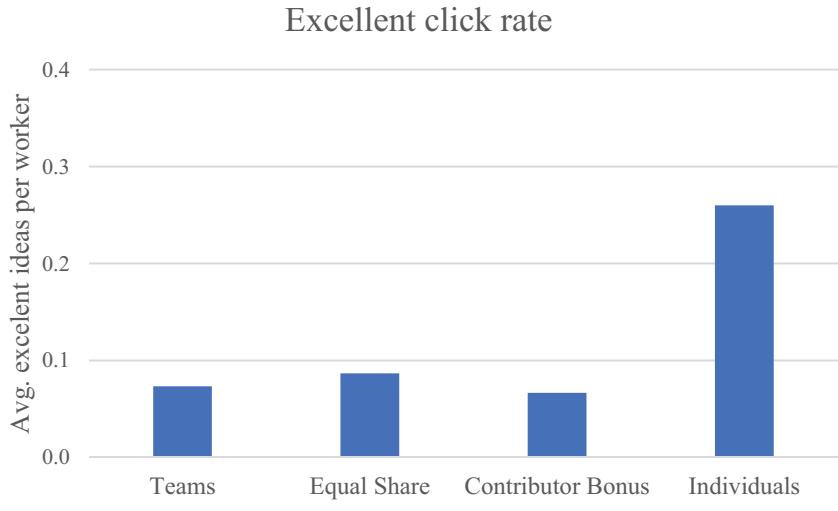

Figure A1. Excellent Ideas per Worker (Excellence in Click Rate).

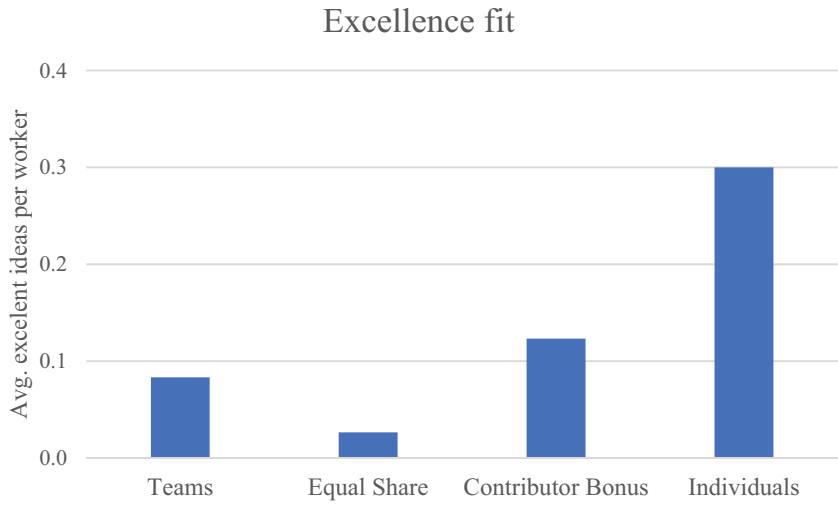

Figure A2. Excellent Ideas per Worker (Excellence in Fit).

The aim is to find a creative title that generates a high click rate and is related to the video. After watching the video, you will have the possibility to communicate with your group members via chat. Your chat name will be Participant X.

Your group will be randomly matched with another group of three participants from this experiment. We will post the title to people on the internet, and observe which title generates more clicks by people who need to watch one of them. The group that will generate the title with the higher click rate will win.

You will have 15 minutes to chat with your group members and to come up with a title. Within these 15 minutes, one group member has to enter the title in the field 'Please enter a title' and the other two group members have to agree to this title. Any of the group members can type in a title at any time in the experiment; afterwards the other two group members will be asked whether they agree to it. A title is only submitted if all group members agree. If your group has not submitted a title before the time is over, you will not receive payment for performing this task.

–next screen –

Your Payment

[Teams and Individuals] In addition to the $5 show-up fee, you will be paid $4 for performing this task.

[Equal share] In addition to the $5 show-up fee, the winning group will earn a $24 prize. This prize will be equally shared between all group members, leaving each member of your group with $8 for performing this task.

[Contributor bonus] In addition to the $5 show-up fee, the winning group will earn a $24 prize. This prize is shared as follows: The group member that entered the final title will earn $14, while the other two members of the winning group (who agreed on this title) will earn $5 each.

–next screen–