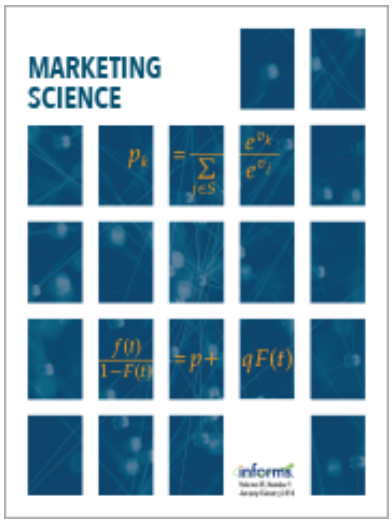


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Raw Ideas in the Fuzzy Front End: Verbosity Increases Perceived Creativity

Laura J. Kornish, Sharaya M. Jones

To cite this article:
Laura J. Kornish, Sharaya M. Jones (2021) Raw Ideas in the Fuzzy Front End: Verbosity Increases Perceived Creativity. Marketing Science

Published online in Articles in Advance 05 Oct 2021

. https://doi.org/10.1287/mksc.2021.1300

http://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE
Articles in Advance, pp. 1–17
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Raw Ideas in the Fuzzy Front End: Verbosity Increases Perceived Creativity

Laura J. Kornish,[a] Sharaya M. Jones[b]

[a] Leeds School of Business, University of Colorado Boulder, Boulder, Colorado 80309; [b] School of Business, George Mason University, Fairfax, Virginia 22030

Contact: kornish@colorado.edu, https://orcid.org/0000-0001-8227-3595 (LJK); sjones72@gmu.edu, https://orcid.org/0000-0001-9583-3379 (SMJ)

Received: February 18, 2019
Revised: November 3, 2019; June 10, 2020; December 28, 2020; April 10, 2021
Accepted: April 21, 2021
Published Online in Articles in Advance: October 5, 2021

https://doi.org/10.1287/mksc.2021.1300

Abstract. At the "fuzzy front end" of an innovation process, organizations typically consider dozens, or even hundreds, of raw ideas. Selecting the best ones is a double challenge: evaluating so many ideas is a large undertaking, and the ideas in their raw form permit only noisy evaluations. In this paper, we demonstrate a further challenge to that large-scale evaluation of raw ideas. We show that verbosity raises the evaluation of ideas, that is, ideas expressed in more words are rated higher. This relationship is especially pronounced for ratings of creativity. Theory tells us that the effect of length on creativity is compounded because length cues both components of creativity—novelty and usefulness. We demonstrate how effort in reading (disfluency) and perceptions of complexity work together to explain the relationship between length and creativity. Our findings provide simple but important new directives for improving the use of crowdsourcing in the practice and study of innovation: either standardize the length of the ideas or control for length in their evaluation. Overall, we urge care with using measures of novelty or creativity when the idea descriptions vary in length.

History: Olivier Toubia served as the senior editor for this article.
Supplemental Material: Data and the online appendices are available at https://doi.org/10.1287/mksc.2021.1300.

Keywords: innovation • idea generation • idea selection • creativity • new product development • crowdsourcing • fuzzy front end

At the "fuzzy front end" of an innovation process, organizations typically consider dozens, or even hundreds, of raw ideas. Sorting through a large set of ideas to identify the most promising ones is a significant challenge. Academics have tackled this challenge from various angles. There are studies that examine the success of an idea based on its origin (Goldenberg et al. 2001, Girotra et al. 2010, Poetz and Schreier 2012, Rosenzweig et al. 2015), its semantic attributes (Eliashberg et al. 2007, Toubia and Netzer 2017), its more general attributes (Åstebro and Elhedhli 2006, Scott et al. 2020), and its fundamental structure (Goldenberg et al. 1999).

Like the authors of those papers, we are also concerned with insights for identifying promising ideas. In this paper, we tackle an issue that occurs when diverse crowds of people contribute ideas, increasingly common with the use of crowdsourcing in innovation. The issue is that the ideas include different levels of elaboration, and what may seem like cosmetic differences have predictable implications for the evaluations of the ideas.

We present evidence for a robust pattern, that people rate ideas expressed with longer verbal descriptions more favorably, especially for ratings of creativity. Knowledge of this pattern is both useful and concerning. This knowledge is useful because we can attempt to account for the bias toward length. But this knowledge is concerning because when crowds of people generate ideas, there is great variation in the length of the ideas. A longer expression does not necessarily fundamentally improve or even change the idea, but the longer version will tend to be rated as more creative.

The pattern that longer ideas are rated as more creative shows in observational data and in our experiments. We also find some evidence for a quadratic relationship between length and ratings of creativity. The quadratic term reflects a decreasing marginal effect of length on ratings. In the large sets of ideas that we analyze, we see very few ideas in which excessive length seems to have a net negative effect on perceived creativity.

There are several plausible explanations for why longer ideas are perceived as more creative. Creative ideas are those that are both useful and novel. Previous authors have found product descriptions that

mention more features have higher purchase intention and perceived effectiveness (Carpenter et al. 1994, Brown and Carpenter 2000, Thompson et al. 2005, Sela and Berger 2012). More fundamentally, longer arguments are considered more persuasive under peripheral processing (Petty and Cacioppo 1984). These patterns plausibly explain the connection between length and usefulness.

There are also multiple reasons that longer idea descriptions could be perceived as more novel. Longer ideas can have—or appear to have—greater complexity. Complexity is defined as "composed of many interconnected parts," and longer descriptions have more words, which are the "parts" of a description. Each additional part is a candidate to make the idea seem novel. Longer ideas can also take more effort to read, reducing processing fluency and a feeling of familiarity (Schwarz 2004, Morewedge and Kahneman 2010, Schwarz 2010). We collect data to sort out whether complexity and disfluency are alternative or complementary explanations.

In weighing evidence for the plausible pathways, we find the strongest evidence for this explanation: longer ideas are perceived as more creative because they appear more complex. Disfluency plays a supporting role, not increasing creativity directly, but indirectly, by raising perceptions of complexity. Therefore, two consequences of length—perceived complexity and disfluency—act together to increase perceived creativity. Contrary to our intuition (and perhaps yours), these effects of length do not appear to be due to deficits of attention. Rather, the added length, even when it doesn't change the substance of the idea, differentiates the ideas in the way Carpenter et al. (1994) and Brown and Carpenter (2000) predict.

We want to be clear that this paper addresses perceived creativity of ideas. We are not claiming that making an idea description longer makes the underlying idea more creative in some intrinsic sense. But the effect on perception is important because assessments of raw ideas in the fuzzy front end drive decisions about which ideas will receive further investment.

The paper is organized as follows. First, we explain the theoretical underpinnings of our investigation, elaborating on why the length of the idea description influences evaluations. Second, we present observational evidence from nine sets of ideas to show that longer ideas are rated as more creative. Third, we present experimental evidence for the causal effect of length on evaluations. Fourth, we rule out an alternative explanation from a recent result in the marketing literature. Fifth, we explore which of the possible theoretical explanations for the phenomenon has the most support. Sixth, we propose remedies for this systematic pattern. Finally, we conclude.

Why Longer Descriptions Are Perceived as More Creative

Why would people perceive longer idea descriptions as more creative? There is a strong consensus that creative ideas are ones that are both useful and novel. (Runco and Jaeger (2012) give a full history of this consensus.) We argue that the description length increases perceptions of both usefulness and novelty, resulting in a strong effect that length increases perceived creativity.

Usefulness

A stream of literature in marketing shows that longer descriptions of products through inclusion of more attributes improves the overall evaluations of the products. The improved evaluations include greater perceptions of capability and higher likelihood of choice from a set of options. Aaker (1991, p. 97) previews these findings by describing a general more-is-better reasoning. He explains that "seemingly trivial but observable" attributes related to quantity or volume improve evaluations of capability; bigger speakers give better sound; more suds clean better.

Carpenter et al. (1994) find that additional attributes in product descriptions—even if those attributes are meaningless—still provide a basis for differentiation from competitive options. Brown and Carpenter (2000) build on that work by identifying conditions under which mentions of "trivial attributes" improve evaluations of products. Their central argument is that the trivial attributes can be instrumental in a task goal; additional attributes can make a choice, or a rating, easier by providing a reason for the choice (Shafir et al. 1993). Thompson et al. (2005) advanced the discussion from overall evaluations to more specific estimates of the capabilities of a product. They find that mentioning more features improves estimates of capability of products. More recently, Sela and Berger (2012) replicate the result that the number of attributes in a product description affects perceptions of capability and usefulness. They show the effect for both hedonic and utilitarian options (and their key finding is that there is a bigger boost in evaluations of the hedonic ones). Thus, there is existing evidence that more verbose descriptions are received more positively in general, and more specifically, that they raise perceptions of the capability or usefulness of the ideas described.

These findings about the effect of added description length are consistent with well-documented heuristics related to amount of information. An early and well-known study about the positive effect of description

length is set in a persuasion context. Petty and Cacioppo (1984) hypothesized that the number of arguments, a variable akin to description length, elicits a more-is-better heuristic. They found that when subjects weren't highly invested in the topic of an argument, they processed heuristically, and length was persuasive. Specifically, they found that a list of six arguments generated more agreement than a list of three. A related idea to the length-persuasion connection is that frequency of information also serves as a positive signal. Alba et al. (1994) analyze "the frequency heuristic" in a price advertising setting. They explain that the count of a cue is easily encoded and serves as a positive signal.

Reading and rating descriptions of raw ideas in the fuzzy front end of innovation is a naturally low-involvement activity that encourages heuristic processing. There are a lot of descriptions and some of them contain just the germ of the idea. The ideas don't have a lot of information to dig into, even if raters wanted to systematically process. The repetition involved in evaluating many ideas also discourages reflection. The repetitive setting matches well with how Chaiken and Maheswaran (1994, p. 460) describe heuristic processing, "any attitude formation or change mechanism that causes persuasion in the absence of argument scrutiny."

These multiple, converging effects—influence of additional, even meaningless, attributes on evaluations and information-quantity heuristic cues—together support the hypothesis that longer ideas will be perceived as more useful. That longer ideas are considered more useful is one key piece of the argument for why longer ideas would be considered more creative.

Novelty

The other key piece of the argument for why longer ideas seem more creative is the relationship between length and novelty. Again, the literature offers multiple pathways for this relationship. We discuss the roles of perceived complexity and processing fluency on perceptions of novelty.

The concept of complexity shows up in studies of physical, biological, human, and other systems (Kauffman 1984, Levinthal 1997, Page 2015). Across the disciplines, the concept has a foundation in interconnection of the parts, reflected in the definition of the word complex, "composed of many interconnected parts." In our setting, ideas with longer descriptions have more parts (words), and therefore appear to be more complex. Other perspectives support the same conclusion, that description length increases complexity. Lurie (2004) analyzes information overload that comes with long descriptions. And Broniarczyk and Griffin (2014) show evidence that more information increases the task complexity.

The literature does not contain a definitive causal pathway from complexity to novelty, but there is strong evidence of an association in marketing scholarship. Hoeffler's (2003) work focuses on how uncertainty affects perceptions of novelty. He cites the seminal work of Rogers (1995) in pointing out that complexity can be a source of uncertainty. Other authors examine the interplay of the two concepts. Mukherjee and Hoyer (2001) show an interaction between novelty and complexity: how people think about novel attributes differs based on the level of the product's complexity. Mugge and Dahl (2013) also account for the role of complexity in perceived novelty.

Older work examines more fundamental ties between complexity and novelty. Berlyne (1960, 1970) underscores the importance of complexity and novelty in stimulus selection. Stang (1977, pp. 319–320) concludes that complexity and novelty are psychologically equivalent. He explains that "Simple stimuli are likely to be reminiscent of many other things, thus seeming familiar," whereas "complex stimuli are likely to be reminiscent of little, thus seeming novel." Although the empirical work leading to these conclusions is based on visual, as opposed to verbal, stimuli, his logical arguments apply to both modalities.

Processing fluency is another plausible pathway from length to novelty. Results on processing fluency tell us that when a person exerts more effort to process a stimulus, the difficulty has predictable effects on perceptions of the stimulus (Schwarz 2004, 2010; Morewedge and Kahneman 2010). The ease, or fluency, has many effects, and an important one is that the stimulus seems more familiar. Harder-to-read passages seem less familiar. In other words, the more work it is to read a description, the less familiar and more novel it will seem. In an objective sense, longer ideas are more work to read than shorter ones; there are more words to read, and it takes more time and effort to perform the task.

Both complexity and processing fluency can explain the link from length to novelty. They are not mutually exclusive, and the direction of the effects could go either way. It could be that longer ideas are perceived as more complex, and that perception of complexity makes something more effortful to process (less fluent), making it seems less familiar. Or, because longer ideas are objectively more work to read, that makes them processed less fluently and therefore perceived as more complex. We test these pathways in our analyses to see which has the strongest empirical support.

Summary and Other Considerations

Returning to our core questions, why would people perceive longer idea descriptions as more creative? Our overarching explanation is that length increases perceptions of both usefulness and novelty, the two

main ingredients of creativity. In our studies, we probe which of the explanations from the literature for each pathway have the best empirical support.

One pattern we see that supports the overall conceptual model is the differential effect of length on ratings of creativity and purchase intent. The two effects of length, on usefulness and on novelty, each enhance perceptions of creativity. The same convergence is not generally true for purchase intention. Perceived usefulness does increase intention, but novelty has been shown to decrease it (Rogers 1995, Hoeffler 2003, Alexander et al. 2008).

As we present our empirical results, we address nuances from alternative consequences of the elements of our model. For example, if disfluency is at work to make longer ideas appear more novel, does it also make them less liked overall? (Reber et al. 1998, Morewedge and Kahneman 2010). We layer on the discussion of complications as we discuss the results from the studies.

Study 1: Observational Evidence That Length Predicts Ratings of Creativity

We start with observational evidence for the relationship between idea description length and perceived creativity. In the subsequent sections, we present results from controlled experiments to establish causation.

Study 1 includes nine data sets. Across the nine data sets, we see a positive and highly significant relationship between idea description length (in characters) and rated creativity. For comparison, we also examine the relationship between length and purchase intent. That relationship is also positive, but weaker. In this section, we describe the data sets and present the observational evidence for the relationships.

TN Data Sets

Seven of the data sets come from Toubia and Netzer (2017). The metrics from these data sets are available to researchers from *Marketing Science*. The topics of these sets include oral care, health-related smartphone apps, and insurance.

Frosting Ideas

The frosting ideas set contains ideas for new frosting products (frosting in the sense of icing typically used to cover cakes and cookies). A panel of creative consumers who contract with a new products consulting firm generated the ideas in response to this prompt:

> For this assignment, you are a brand manager for a consumer products company that has a successful line of frosting products. The frostings come in different flavors and are primarily used for covering cakes, cupcakes, and cookies. The products can be found in grocery store chains, usually in the baking aisle. Your task is to come up with new frosting products, that don't currently exist, that you would sell in the same grocery stores.

There were 248 ideas in total. We collected creativity and purchase intent ratings from Mechanical Turk workers. On average, 31.6 people rated each of the 248 ideas. The range was 18 to 48 raters for each idea. In all studies, we include only respondents who passed an attention filter.

Household Products Ideas

The household products data set is based on one of the idea sets from Kornish and Ulrich (2014, p. 18), ideas submitted to Quirky.com, a community product development website. Quirky's community aimed to develop "consumer products that could retail for under $150 and don't involve integrated software." We obtained the data set from the authors. The data set has 100 ideas randomly selected from the entire universe of thousands of ideas submitted to Quirky. The purchase intent responses were from a general population panel purchased from the vendor Qualtrics as part of the original data set. Each idea was rated for purchase intent by between 282 to 293 people. We independently collected creativity ratings from Mechanical Turk workers. On average, 54.5 people rated each of the 100 ideas for creativity. The range was 51 to 58 raters for each idea.

Results

In all nine data sets, we observe that longer descriptions are significantly positively correlated with higher creativity ratings. See Table 1. The relationship between length and purchase intent is also positive, but weaker. In about half of the data sets, length and purchase intent are significantly positively correlated, and the overall relationship from a meta-analysis of these nine studies between length and purchase intent is statistically significant.

Using the R package *meta*, the function *metacor* reports a 0.34 overall correlation between length and creativity for a random effects model ($p < 0.0001$) and a 0.14 correlation between length and purchase intent ($p < 0.001$).

In addition to the consistent linear relationship between length and perceived creativity across these data sets, there is also reasonably consistent evidence for a concave, quadratic relationship between length and perceived creativity. In six of the nine data sets, adding a quadratic term for length adds significant explanatory power. In all nine of the data sets, the quadratic term has a negative coefficient, and the linear term maintains a positive coefficient. These patterns imply that in these data sets, on average, longer ideas

Table 1. Correlations with Idea Description Length in Nine Studies (Study 1)

Number of ideas in the data set and data set topic (TN indicates Toubia and Netzer 2017)	Length in number of characters mean (SD)	Correlation of length and creativity	Correlation of length and purchase intent
248 frosting products	124.7 (134.4)	0.31***	0.16*
100 household products	745.4 (470.1)	0.31**	0.05
276 new insurance products related to aging (TN study 1a)	389.5 (210.3)	0.34***	0.10
271 new insurance products related to financial security (TN study 1b)	356.8 (197.9)	0.44***	0.12+
251 new insurance products related to unemployment (TN study 1c)	412.3 (193.1)	0.40***	0.12+
555 health-related smartphone apps (TN study 2)	215.1 (189.2)	0.14***	0.15***
173 health-related smartphone apps (TN study 3)	179.7 (140.6)	0.44***	0.21**
220 oral care for women over 40 (TN study 4)	Not available	0.16*	−0.04
648 health-related smartphone apps (TN study 6)	Not available	0.46***	0.28***

Note. Not available indicates that we did not have the text of ideas in that study.
$+p < 0.1$, $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

are considered more creative, and there is some evidence to a decreasing marginal return from length. Only a small percentage, if any, of the ideas in each data set is long enough to be past the peak of the estimated quadratic function. Online Appendix A contains details of that analysis.

With observational evidence, we can't conclude whether longer ideas are perceived as more creative or ideas that are more creative take more words to express. We next present the results of an experiment to provide evidence for causation.

Study 2: Does Length Improve Evaluations? Experimental Evidence

To establish evidence for a causal relationship—that making an idea longer increases the perception of creativity—we ran experiments manipulating the length of ideas.

In Study 2, we separate the effects of length and the inherent merit of the idea by evaluating three different versions of ideas. We used an original version, a lengthened version that added no information, and a lengthened version that added concrete but unsurprising details. We test to see if the lengthened versions of the same idea are rated higher on creativity than the original version. We expect that the lengthened idea with the details would rate higher than the original, based on our analysis of the nine observational studies. This study also allows us to see whether the lengthened-but-no-more-informative version is rated more like the original version or more like the lengthened version with details.

Stimuli

In this study, we used a set of ideas for products for the college student market. We randomly selected 29 ideas from a larger set that students generated for a course project for "physical products for the college student market with retail price of under $50."

We worked with a subset of the ideas (only 29 rather than the full set, which had 290) for a few reasons beyond the obvious one of frugality. First, we have found that our respondent pool attends well to a task of reading and rating about 30 ideas. When the set of ideas is higher, closer to 50, attention filter pass rates drop off. Second, when all respondents rate the same set of ideas, we can look at consistency of the ratings across raters. Third, with a sample size of 29, we will only detect effects of at least moderate strength.

In this study, we created two variations on each idea, both longer than the original idea. One of the longer versions (version B) was extended by repeating information in the original idea (version A) and adding only obvious statements. The other longer version (version C) included additional concrete details, but ones that were straightforward elaborations on the original idea. The longer versions were 1.4 to 2.3 times as long as the original version.

The following are three versions of one of the ideas:

• Version A: This product involves using durable plastic to replicate a large sized cardboard box, with in-built slots that allow shelves to create separate compartments within. This will solve the problem of storing fragile items with everything else. (239 characters)

- Version B: This product involves using durable plastic to replicate a large sized cardboard box, with in-built slots that allow shelves to create separate compartments within. This box will be made of plastic so it is durable and it will be pretty large. This compartment box will solve the problem of storing fragile items with everything else because the separate compartments will keep fragile items in place. (403 characters, 1.7 times as long as version A)
- Version C: This product involves using durable plastic to replicate a large sized cardboard box, with in-built slots that allow shelves to create separate compartments within. The compartments will have removable pads that line them. This will solve the problem of storing fragile items with everything else. The box should also have a handle with an ergonomically designed grip to make transporting the box easy. (403 characters, same as version B, adding the unsurprising but concrete details about the pads and handle)

Manipulation Check

First, we checked that our construction of the three versions was perceived as intended. For each of the 29 ideas, we asked a small set of people to compare either versions A and B of the 29 ideas or versions A and C. The people were randomly assigned to one of the conditions. For both conditions, the question was as follows. "The two versions describe the same idea, but the descriptions are not exactly the same. Considering the content of the descriptions, how big a DIFFERENCE do you see in the two versions?" The response options were (1) no or essentially no difference, (2) a small difference, (3) a moderate difference, and (4) a large difference. Our manipulation was confirmed. The average rating of the six people who compared versions A and B was 1.83 (i.e., between no difference and a small difference). The average rating of the three people who compared versions A and C was 2.31 (i.e., between a small difference and a moderate difference). The paired t-test for the 29 ideas gives $t = 4.79$ ($p < 0.0001$). This confirms our manipulation in two ways: for both comparisons, the differences are not perceived as large; and the difference between versions A and B is smaller than the difference between versions A and C.

Idea Ratings

Next we collected creativity, market uniqueness ("How different is this idea from products on the market today?"), and purchase intent ratings of all three versions of the 29 ideas from Mechanical Turk workers. Each worker was randomly assigned to rate one block of 29 ideas, including exactly one of the three versions for each idea, with the versions interleaved in the blocks (i.e., block 1 had the pattern A, B, C, A, etc.; block 2 had the pattern B, C, A, B, etc.; and block 3 had the pattern C, A, B, C, etc.) We used the blocks to ensure that each person saw a balance of the ideas and the conditions. We received 50–52 responses that passed the attention filter in each block.

Results

Table 2 shows the differences in the means on the three measures—creativity, market uniqueness, and purchase intent. The original, version A, is rated as less creative than the longer versions. The differences for market uniqueness show the same signs as those for creativity. We observe essentially no difference (and definitely not a statistically significant difference) in purchase intent across the three versions.

Discussion

Controlling for the content of the idea by using the paired structure of the stimuli, we find that longer descriptions are perceived as more creative. Ideas that are expressed with more words, even if there is no additional important information in the longer version, are considered more creative. Both versions of the longer ideas are rated higher on creativity than the shorter version. This result is consistent with length providing cues for creativity.

Online Appendix B contains a variation on the analysis standardizing each person's set of responses to adjust for people's differential use of the scale. The results are very similar. The most notable difference is that the difference in market uniqueness between versions A and B is significant in the scaled analysis.

Although the differences in creativity and uniqueness between versions B and C are not significant, they are similar in magnitude to the differences between A and B. This pattern is suggestive that there is some effect of the concrete details in version C. Tversky's (1977) features of similarity model predicts that people will consider objects with more distinctive features to be more novel, holding similar features constant. However, the significant difference between versions A (original) and B (extended) shows that concrete details are not necessary to raise the perception of creativity.

In the observational data, we did see an overall association between length and purchase intention, but it was weaker than the association of length and creativity. In the next study, we apply a stronger manipulation of length and expect more of an effect on purchase intention.

Study 3: Does Length Improve Evaluations? Replication

We ran another study with a similar structure to Study 2 to examine the effect of the length of idea descriptions on the ratings of the ideas. This study used

Table 2. Differences in Evaluations for Versions of Ideas in Study 2

Version	Creativity	Market uniqueness	Purchase intent
Version A (original) − B (extended)	−0.10*	−0.07	0.00
	(0.04)	(0.05)	(0.06)
	$p = 0.04$	$p = 0.12$	$p = 0.97$
Version A (original) − C (extended with detail)	−0.17*	−0.15*	0.00
	(0.06)	(0.06)	(0.07)
	$p = 0.01$	$p = 0.02$	$p = 0.99$
Version B (extended) − C (extended with detail)	−0.08	−0.07	0.00
	(0.06)	(0.06)	(0.07)
	$p = 0.23$	$p = 0.20$	$p = 0.98$

Notes. Each cell shows the difference in means, the standard error of differences, and the *p*-value for a paired t-test on 29 pairs. The unit of analysis is the idea version. This data set has 29 ideas, three versions of each.
*$p < 0.05$.

a different random subset of 29 ideas from the 290 ideas for college student products, and we created two other versions of each idea. In Study 3, the three versions of the ideas were the original idea, a very short version with only a few words, and a lengthened but essentially the same idea version. This study has a much heavier manipulation of length, in both directions, than Study 2. We expect a more pronounced effect from this stronger manipulation.

Stimuli

In this study, we used repetition, use cases, unsurprising details, and options to add length. The shortened version is brutally short at around 30 characters. The short version omits many details, retaining only a noun phrase with a word or two describing the object (bottle opener, drink coaster, drinking glass) and a few words offering the key distinguishing feature.

We used the same blocked and interleaved design as Study 2 so raters saw only one version of each idea. We collected ratings from 54 to 61 Mechanical Turk workers (those passing the attention filter) in each block. We collected ratings on four variables: creativity, market uniqueness, purchase intent, and need. Table 3 shows the results.

Results

There is a clear positive association between length and creativity. The longer version (B) was most creative with a mean of 3.47, then the original version (A) with a mean of 3.30, and then the shorter version (C) with a mean of 2.70. This study replicates the pattern that longer descriptions are considered more creative. As in Study 2, the results for uniqueness are a less pronounced version of those for creativity, with the shorter version rated as less unique than the others, but only a small (and not statistically significant) difference between the original and longer versions. Online Appendix B contains a variation on the analysis standardizing each person's set of responses to adjust for people's differential use of the scale. The results are very similar.

For purchase intent, Study 3 does show an effect of length; Study 2 did not. In the meta-analysis of our observational data (Study 1), we did see a significant effect of length on purchase intent. The different pattern for purchase intent and need between Study 2 and Study 3 is consistent with our conclusion that the effect of length on purchase intent is weaker than the effect of length on creativity. In Study 3, there is a more drastic difference in length between the versions compared with Study 2. In Study 3, the long versions

Table 3. Differences in Evaluations for Versions of Ideas in Study 3

Version	Creativity	Market uniqueness	Purchase intent	Need
Version A (original) − B (longer)	−0.17**	−0.06	−0.14**	−0.14***
	(0.05)	(0.04)	(0.05)	(0.03)
	$p = 0.002$	$p = 0.14$	$p = 0.007$	$p = 0.000$
Version A (original) − C (shorter)	0.59***	0.40***	0.19***	0.10*
	(0.07)	(0.08)	(0.05)	(0.05)
	$p = 0.000$	$p = 0.000$	$p = 0.001$	$p = 0.04$
Version B (longer) − C (shorter)	0.76***	0.46***	0.32***	0.24***
	(0.08)	(0.08)	(0.07)	(0.05)
	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$

Notes. Each cell shows the difference in means, the standard error of differences, and the *p*-value for a paired t-test on 29 pairs. The unit of analysis is the idea version. This data set has 29 ideas, three versions of each.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

were 2.0 to 4.8 times as long as the originals; that range was 1.4 to 2.3 for Study 2. And in Study 3, the short versions were very short. The short versions were 0.12 to 0.24 times as long as the originals. Study 3 uses bigger differences in length and bigger differences in the details included in each version. As predicted, we see bigger differences in the metrics in Table 3 compared with Table 2—big enough compared with the variability to be statistically significant.

Study 4: Length and Distance from Prototypicality

Having established a strong positive relationship between idea description length and perceived creativity, we now examine a possible alternative explanation. Work by Toubia and Netzer (2017) (which we refer to as TN) studies the relationship between perceived creativity of ideas and their semantic network properties. We present evidence that our result, that length increases perceived creativity, is not explained by TN's result about the balance of close and far associations in the idea.

The work of TN contains a powerful and general technique for predicting perceived creativity of ideas and suggesting ways to increase it. They use a semantic network approach to analyze raw ideas, and their focus is proof that a prototypical balance of close and far associations between pairs of words predicts creativity. They measure distance from the prototypical balance by the maximum distance between two distributions. One distribution is for the edge weights in the semantic network of a single idea (where the words in the idea are nodes). The other distribution is the average of the edge weight distributions over some reference corpus (e.g., text of webpages from a related Google search or a set of related ideas).

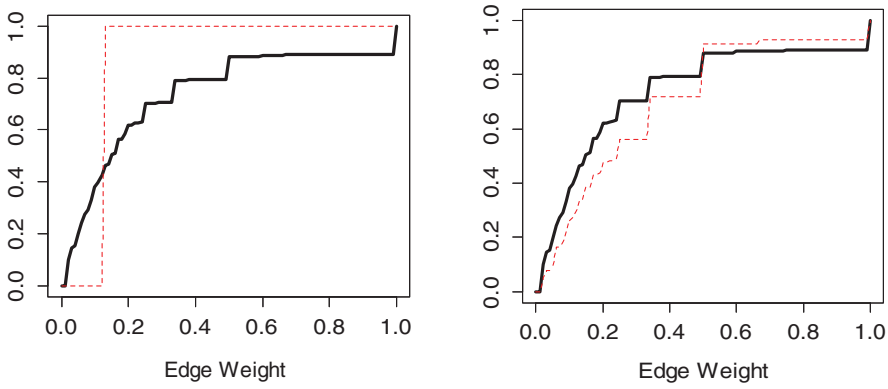
Figure 1 illustrates the distance measure. The figure shows edge weight distributions for two ideas from the frosting ideas data set we introduced in Study 1. Both graphs in the figure show the prototypical cumulative distribution function of edge weights as a solid black line: the vertical value is the fraction of edge weights that are less than or equal to each horizontal axis value. The edge weights are between 0 and 1, with 1 meaning perfect association (the word on one side of the edge never appears without the word on the other side of the edge). Each graph also has a dashed line, representing the edge weight distribution for a single idea. The idea on the left is "Cookie dough" and the idea on the right is "Frosting with the ability to be easily drizzled over a cake for a ganache coating. Maybe it needs to be heated up in some way or is already pourable and packaged to pour over a cake and harden slightly in the air temperature (think like Magic Shell hardens over ice cream)."

The network for the idea "Cookie dough" has only two nodes, and therefore one edge. The single edge creates the single-stepped shape of the dashed-line cumulative distribution function in the left graph. The network for the other, longer idea has many nodes and therefore many edges. Its cumulative distribution has more steps and the steps are smaller.

TN note the confounding of length and distance: "larger semantic subnetworks tend to have smoother distributions of edge weights, which tend to be more prototypical" (Toubia and Netzer 2017, p. 8). To statistically account for this potential confound, they include idea description length and semantic network size of the idea as control variables in their regressions. TN's table 2 shows that, after controlling for distance, length is positively related to creativity in six out of the eight studies and significantly so in five.

Considering their findings in the context of our results, it is natural to ask whether the length effect is distinct from, or is it subsumed in, the distance from prototypicality effect? Is distance from prototypicality an alternative explanation for the effect of length on

Figure 1. (Color online) Cumulative Distribution Functions for Two Frosting Ideas



Note. In this example, the baseline semantic network is created from the ideas themselves.

perceived creativity? To answer that question, we run an experiment that disentangles the effects of idea description length ("length") and distance of an idea's semantic subnetwork from the prototypical edge weight distribution ("distance"). We independently manipulate these two factors and show that the effect of length is above and beyond that of distance.

Stimuli

To generate the stimuli for this study, we obtained the ideas generated for TN studies 2 and 3, "ways your smartphone can make you healthier," from the authors. We randomly selected 25 ideas out of the full set ideas from the two studies. (Table 1 shows 555 + 173 = 728 ideas in the two studies. The full set from the authors had 752 ideas.) We then crafted four versions of each of those 25 ideas: one with a relatively low distance and short description, one with a relatively high distance and short description, one with a relatively low distance and long description, and one with a relatively high distance and long description.

To generate the variations, we initially wrote 10 versions of each idea, some shorter and some longer. Then we calculated length and distance for each of the 10 versions.[1] Within the 10 variations, we looked for a set of four that were fully crossed on length and distance. If we didn't have a fully crossed set within the first 10, we wrote additional variations of the descriptions in a more targeted manner, for example, lengthening or shortening descriptions to fill a missing cell, or examining the stems in the semantic subnetworks of the idea versions to craft a version with a particular distance. We acknowledge that as the authors, we are not blind to the research questions. We made every effort to write ideas equivalent in content, just varying in length and distance. The full set of stimuli appear in Online Appendix B.

Within the four versions of each idea, we ensured that length and distance were orthogonal. The two short versions were the same length (within five characters), as were the two long versions. The two high distance versions were the same distance (within 0.03, recalling the distance measure ranges from 0 to 1), as were the two low distance versions. Further, we wanted noticeable separation between the long and short and high and low versions within each pair. The differences of high versus low distances range from 0.128 to 0.500, with an average of 0.246 across the 25 ideas. The differences in long versus short lengths range from 67 characters 188 characters, with an average of 124.7 across the 25 ideas. (The ratios of the distances of long to short range from 1.70 to 3.41, averaging 2.31.) This structure disentangles length and distance.

We also ensured that the distributions on length and distance reasonably followed the benchmark of the full set of original ideas in TN's studies 2 and 3. Where there was a deviation, we ensure there was a little more variance in distance and less variance in the length in our stimuli, compared with the original set of ideas. The interquartile range (difference of 75th and 25th percentiles) for distance is 0.13 for the original ideas and 0.20 for our stimuli. The interquartile range for length is 172 characters in the original ideas and 137 in our stimuli. Similar comparisons hold for other ranges of percentiles. These deviations make our test of our proposed effect of length a conservative test, that is, likely to understate the effects.

We used the same approach to creating blocks of ideas as in Study 2 and Study 3, and we hired Mechanical Turk workers to rate of the ideas. Each worker saw one version of each idea, with the versions and ideas balanced across the conditions. Each worker rated creativity and purchase intent for the 25 ideas in their assigned block. Each block received between 52 and 57 responses that passed the attention filter.

Results

We analyzed this data using a mixed regression model, with fixed effects for the two factors and the interaction and random effects for the idea. We used the values for distance and length in the analysis, not the binary indicators of long versus short (length) and high versus low (distance). A mixed regression model allows us to examine both factors (collapsing across conditions) while controlling for variation by idea. Table 4 shows the coefficients for the standardized fixed effects.

The positive coefficient on idea length is clearly significant, with $t = 6.581$: longer ideas are rated as more creative. The coefficient on distance is not significant, although it is negative (consistent with TN) and it approaches significance, with $t = -1.522$. The interaction is not significant.

Discussion

Our experiment shows that people rate longer versions of ideas as more creative, compared with shorter ones. The effect of length on perceived creativity is

Table 4. Estimates for the Standardized Coefficients for the Fixed Effects for Study 4

Variable	Estimate	Std. Error	*t* value
(Intercept)	3.269	0.077	42.326
Distance	−0.035	0.023	−1.522
Idea length	0.138	0.021	6.581
Distance × Idea length	−0.005	0.017	−0.269

Note. The table shows the fixed effects for the model, creativity = distance + idea length + distance * idea length + (distance + idea length | idea), with the idea version as the unit of analysis, where $N = 100$, from 25 ideas × four versions of each.

above and beyond the effect of the distance from prototypical close and far associations. When not independently manipulated, length and distance covary, but Study 4 supports our contention that an idea's distance from the prototypical edge weight distribution is not an alternative explanation for the patterns we see.

Study 5: Replication of Study 2 Plus Complexity and Fluency Mediations

We now turn our attention to explaining why length raises perceptions of creativity. In this section, we present Study 5, which shows that complexity mediates the relationship between length and creativity. We also analyze the role that fluency plays, on top of complexity, in the relationship.

Stimuli

Study 5 is a replication of Study 2 using only version A (original) and version B (extended with repetition and obvious statements) interleaved. We use these two conditions as a strong test of the mechanism because the two conditions have only subtle differences.

We collected four items for each idea: creativity, purchase intention, complexity, and fluency. The first three questions are in multiple choice format. The last question uses the fluency item validated by Graf et al. (2018). The question uses a slider for the prompt "The process of reading the idea description above was:" with values from 0 (difficult) to 100 (easy). For reasons we explain later, we presented the ideas in each block in the same random order to all raters. There was no effect of the order. We collected ratings from 56 Mechanical Turk workers (those passing the attention filter) in each block. Table 5 summarizes the results.

Like the results in the original Study 2, version B is more creative than version A, and there is not a significant difference in purchase intention. As predicted, version B is more complex. Version A is easier to read, that is, the shorter ideas are more fluent. Longer descriptions could be more fluent if internal repetition promotes familiarity within a passage. Indeed, Nunes et al. (2015) find that repetition in song lyrics promotes fluency. But our data show that shorter is more fluent.

Mediation for Creativity: Complexity vs. Fluency

We analyze whether complexity mediates length and creativity and whether fluency does. In other words, does the variation in complexity explain the relationship between length and creativity? What about the variation in fluency? Table 6 shows the results. All mediation analyses use the Preacher and Hayes (2004) routines implemented in the PROCESS macro in SPSS (with 5,000 resamples to estimate 95% confidence intervals). This analysis uses their model 4. All the mediations also control for the idea, that is, include a fixed effect for each of the 29 ideas.

These analyses reveal that complexity is a mediator of the relationship between length and creativity. Fluency, however, does not mediate, given the confidence interval on the coefficient for the indirect effect of fluency (−0.22, 0.34) squarely contains 0. The parallel mediation provides additional evidence for complexity as a mediator over and above any influence of fluency. The nonsignificant direct effect for the complexity mediation means an omitted mediator is unlikely (Zhao et al. 2010).

Theoretical accounts suggest further analyses. Complexity could produce disfluency—more complex descriptions are more effort to read—or the reverse—descriptions that are harder to read appear more complex. Our initial mediation results point to the latter order because the path from length to fluency is significant (coefficient −12.76, standard error (SE) 3.91, confidence interval −20.59 to −4.94; not shown in Table 6), and the path from fluency to creativity is not significant (also not shown). For completeness, we present both serial mediations (5,000 resamples, PROCESS model 6; Preacher and Hayes 2004) in Table 7.

As predicted, only one of the serial paths shows significance: length → fluency → complexity → creativity. The positive sign on the coefficient on the indirect path via fluency→ complexity comes from the negative signs on two paths, from length to fluency and from fluency to complexity.

When ideas are longer, they are more effort to read (more disfluent), so they are perceived as more complex. Rather than mediating length and creativity, fluency mediates length and complexity on the path to creativity. Fluency is a complementary—not an alternative—explanation to complexity. Fluency helps

Table 5. Differences in Evaluations for Versions of Ideas in Study 5

Version	Creativity	Purchase intention	Complexity	Fluency
Version A (original) − B (extended)	−0.08*	0.008	−0.10*	1.25**
	(0.03)	(0.04)	(0.04)	(0.44)
	$p = 0.02$	$p = 0.83$	$p = 0.02$	$p = 0.008$

Notes. Each cell shows the difference in means, the standard error of differences, and the *p*-value for a paired t-test on 29 pairs. The unit of analysis is the idea version. This data set has 29 ideas, two versions of each.

*$p < 0.05$; **$p < 0.01$.

Table 6. Mediation Analyses for Creativity in Study 5

X → M → Y	Direct effect X → Y	Indirect effect for complexity	Indirect effect for fluency
Length → complexity → creativity	0.33	0.37*	
	(0.32)	(0.17)	
	$p = 0.31$	CI: (0.07, 0.73)	
Length → fluency → creativity	0.68*		0.015
	(0.33)		(0.14)
	$p = 0.045$		CI: (−0.22, 0.34)
Parallel mediation:	0.40	0.41*	−0.11
Length → (fluency and complexity) → creativity	(0.33)	(0.19)	(0.16)
	$p = 0.23$	CI: (0.08, 0.85)	CI: (−0.44,0.22)

Notes. The table shows the confidence intervals (CIs) on the mediation path (the indirect effect) as well as the direct effect. The evidence for the mediation is that the CI for the indirect effect excludes 0. The unit of analysis is the idea version. This analysis has 29 ideas and two versions of each, so $N = 29 \times 2 = 58$.
*$p < 0.05$.

explain why longer ideas are perceived as more complex and, therefore, creative.

Discussion

Although a statistical mediation analysis does not prove causality, it does provide some evidence for the mechanism. Study 5 resolves the particulars for our theoretical account of the relationship between length and creativity. It shows that complexity mediates the relationship between length and creativity. It further shows that fluency alone does not explain the relationship; it works in conjunction with complexity. Longer ideas take more effort to read, leading to a heightened perception of complexity and associated increases in ratings of creativity. We replicated the complexity mediation and the serial mediation in supplemental studies 5' and 5'', respectively, which we present in Online Appendix B.

Mediation of Length and Purchase Intention

In Study 1, we see that length and creativity have a stronger association than length and purchase intention. In the nine observational studies reported in Study 1 and in our experiments, we see mixed results for the relationship between length and purchase intention, sometimes positive and sometimes null. In this section, we explore that relationship for Study 5.

Purchase intention is weakly (at best) associated with novelty and strongly associated with usefulness. (In supplemental study 5', reported in Online Appendix B, purchase intention and usefulness are correlated $r = 0.80$ and load on the same factor in the factor analysis.) In contrast, creativity is more strongly associated with novelty than with usefulness (Kudrowitz and Wallace 2013, Diedrich et al. 2015, Acar et al. 2017, Berg 2019). Thus, we would not expect the mediation pathways between length and creativity to also hold for length and purchase intention. Table 8 shows the mediation analyses for purchase intention from Study 5, using Preacher and Hayes's (2004) model 4 with 5,000 resamples.

The coefficients on the indirect paths including fluency are −0.25 and −0.29. They are negative because length reduces fluency, and fluency improves purchase intention. The more fluent a description, the more positively readers will evaluate it (Reber et al. 1998, Schwarz 2004, Morewedge and Kahneman 2010, Schwarz 2010). Purchase intention is a valenced evaluation, and we see the positive association between

Table 7. Serial Mediation Analyses for Creativity in Study 5

X → M1 → M2 → Y	Direct effect X → Y	Indirect effect for complexity	Indirect effect for fluency	Indirect serial effect
				Complexity → fluency
Length → complexity	0.40	0.41*	−0.07	−0.04
→ fluency	(0.33)	(0.19)	(0.11)	(0.08)
→ creativity	$p = 0.23$	CI: (0.10, 0.85)	CI: (−0.29, 0.14)	CI: (−0.24, 0.07)
				Fluency → complexity
Length → fluency	0.40	0.28*	−0.11	0.13*
→ complexity	(0.33)	(0.14)	(0.16)	(0.09)
→ creativity	$p = 0.23$	CI: (0.05, 0.61)	CI: (−0.45, 0.21)	CI: (0.008, 0.37)

Notes. The table shows the confidence intervals (CIs) on the mediation path (the indirect effect) as well as the direct effect. The evidence for the mediation is that the CI for the indirect effect excludes 0. The unit of analysis is the idea version. This analysis has 29 ideas and two versions of each, so $N = 29 \times 2 = 58$.
*$p < 0.05$.

Table 8. Mediation Analyses for Purchase Intent in Study 5

X → M → Y	Direct effect X → Y	Indirect effect for complexity	Indirect effect for fluency
Length → complexity → purchase intention	0.0004	0.021	
	(0.32)	(0.16)	
	$p = 0.99$	CI: (−0.34, 0.32)	
Length → fluency → purchase intention	0.27		−0.25*
	(0.30)		(0.16)
	$p = 0.37$		CI: (−0.62, −0.0002)
Parallel mediation:	0.18	0.13	−0.29*
Length → (fluency and complexity)→ purchase intention	(0.32)	(0.17)	(0.18)
	$p = 0.57$	CI: (−0.19, 0.48)	CI: (−0.69, −0.012)

Notes. The table shows the confidence intervals (CIs) on the mediation path (the indirect effect) as well as the direct effects. This analysis has 29 ideas and two versions of each, so $N = 29 \times 2 = 58$.

*$p < 0.05$.

fluency and purchase intention. (We provide the serial mediation analyses in Online Appendix B.)

Our observational and experimental studies show that the link between length and purchase intention is weaker than the link between length and creativity. Study 5 helps us understand that comparison. Length has opposing effects on purchase intention: elaboration can boost perceptions of usefulness ("lots of features!") or, as we show in Study 5, dampen it with disfluency ("hard to process!").

Conceptually, disfluency's negative effect on evaluations could also dampen ratings of creativity because creativity encompasses both novelty and usefulness. However, novelty and usefulness are not equally weighted in perceptions of creativity. Novelty is a much bigger driver of perceived creativity than usefulness is (Kudrowitz and Wallace 2013, Diedrich et al. 2015, Acar et al. 2017, Berg 2019). Length acts differently on novelty and usefulness, and those differences help explain the different results we see for creativity and purchase intention.

Is Low Attention a Key Driver of the Effect?

Both peripheral processing and meaningless differentiation predict the pattern we see, that adding to the length of a description improves evaluations. A key distinction between the two accounts is the role of attention. Peripheral processing happens when attention is low, when little scrutiny is applied. Little scrutiny is likely an understatement for evaluators of ideas in the fuzzy front end of innovation, especially when people are rating long lists of raw ideas. Meaningless differentiation is not a story grounded in attention levels. Brown and Carpenter (2000) are explicitly agnostic: "we do not mean to address the lively debate about whether the consumer is actually conscious of this process or about the extent to which the process is under the consumer's deliberate control" (p. 375). Meaningless differentiation doesn't rule out low attention, but it doesn't depend on it either.

This distinction suggests that measures and manipulations of attention will reveal whether there is a strong case for peripheral processing. For example, if low attention drives the effect, then people who spend more time rating ideas should show the effect less. We performed three analyses—based on measured speed of completion, measured order of ideas, and manipulated number of ideas to rate—and none of them confirms that peripheral processing drives the higher ratings of longer ideas. That set of null results casts doubt on a pure peripheral processing explanation for our finding. We briefly describe the null results here and provide the tables of results of the analyses in Online Appendix B.

Test one, on speed of completion: do people who speed through the task show the rate-longer-as-more-creative effect more than people who spend more time? We did not see clear evidence of that pattern. We analyzed response time data for Study 5, splitting time spent across the median in each block to create equal-sized slow and fast groups. The differences between slow and fast are not statistically significant. The fast group does have a directionally bigger effect of length on creativity, but the difference is not significantly different from the slow group.

Test two, on the order of the ideas: do people who are rating many ideas rely more on the heuristic that long is more creative as they progress through the set? To answer this question, we again analyzed the results from Study 5, where we purposely kept the same random order of ideas for all raters, so we could see if the effect increased as people worked. There was no effect of order on the differences between longer and shorter versions in any of the metrics (creativity, purchase intention, complexity, fluency).

Test three, on number of ideas: do people who rate fewer ideas rely less on the heuristic that long is more creative than people who rate more ideas? To answer this question, we ran Study 6, where we randomly assigned people to rate 10 (nine real and one attention

filter) or 30 (29 real) ideas on creativity. We used version A (original) and version B (longer) of Study 3. The effect showed for both the "rate fewer" and "rate more" groups, and there was no significant difference between the groups. No, people who rate fewer ideas do not show the effect less.

Of course, null results do not definitively show that peripheral processing is irrelevant. With a bigger sample size, maybe we would see significant effects. But the set of null results together suggest that something else must be going on beyond low attention. Carpenter et al. (1994) explain that the consumer wonders why a claim would be there if it had no meaning. ("Why would P&G develop a technology to flake coffee crystals, seek and receive a patent for it, and then spend considerable sums promoting it?" p. 341) Consumers' elevated judgments of more verbose descriptions are not errors due to lack of attention. Rather, they are consistent with Grice's (1975) norms, a fundamental set of shared assumptions about communication. In our context, the key Gricean norm is that of quantity, "Make your contribution as informative as required. . . . Do not make your contribution more informative than is required" (Grice 1975, p. 45). The repetition, obvious statements, and concrete but unsurprising details in our longer versions of ideas take on meaning to the readers.

Study 7: Reducing the Effect of Length on Creativity

How can we reduce the effect of length on perceptions of creativity? In Study 7, we show that blocking ideas by length—so that each rater sees only a narrow range of lengths—reduces the effect of length on perceptions of creativity. This remedy follows from the key argument of Carpenter et al. (1994) that people assign meaning to additional but meaningless phrases in comparisons of product descriptions. Blocking similar-length ideas reduces the potential for length comparisons to influence evaluations.

Stimuli

In Study 7, we collected creativity ratings on the set of 248 ideas for new frosting products described in Study 1. In this collection, instead of randomly serving a set of ideas to a rater (as we did in Study 1, where we randomly showed 30 out of the 248 ideas), we grouped the ideas into eight blocks of 31 ideas based on their length and showed each rater a randomly selected block. In this study, we collected ratings from 32 to 42 Mechanical Turk workers (those passing the attention filter) in each block, with an overall average of 38.4 raters per idea. Each worker rated all the ideas in the block for creativity and answered an attention filter question. The ideas within a block were shown in random order. In Study 1, each idea was rated by 18 to 48 people, with an overall average of 31.6 raters per idea.

Results

Grouping the ideas by length eliminated the relationship between length and perceived creativity across the 248 ideas. In Study 1, the correlation between length

Table 9. Effect of Using Blocks by Length in Evaluation

Variable	Model 1 Dependent variable: *Creativity*	Model 2 Dependent variable: *Creativity*
(Intercept)	3.44***	3.49***
	(0.03)	(0.03)
	$p = 0.000$	$p = 0.000$
Idea length (Std)	0.17***	0.30***
	(0.03)	(0.04)
	$p = 0.000$	$p = 0.000$
Idea length (Std) squared		−0.04***
		(0.01)
		$p = 0.000$
Condition (mixed=0, blocked=1)	0.18***	0.13**
	(0.04)	(0.04)
	$p = 0.000$	$p = 0.002$
Condition × length (Std)	−0.16***	−0.30***
	(0.04)	(0.06)
	$p = 0.000$	$p = 0.000$
Condition × length (Std) squared		0.05**
		(0.02)
		$p = 0.002$
N	496	496

Notes. The table shows that there is a significant effect of the condition of data collection: ideas shown in random sets (the mixed condition, coded as 0) versus ideas shown in sets of relatively uniform length (the blocked condition, coded as 1). Idea lengths are standardized (to mean of 0, standard deviation of 1).

**$p < 0.01$; ***$p < 0.001$.

and perceived creativity in this set of ideas is 0.31 ($p < 0.0001$). In Study 7, the correlation is 0.03. For $n = 248$, 0.03 is not significantly different from 0 ($p = 0.58$), and it is significantly different from 0.31 ($p < 0.01$).

Recognizing the quadratic relationship between length and perceived creativity, we check for a moderating effect of data collection condition (randomly mixed as in Study 1 versus blocked by length in Study 7) in a model that includes a quadratic term for length. We find the expected result: blocking by length reduces the relationship between length and creativity compared with the randomly mixed display of ideas. That result shows in Table 9 as the significant interaction terms, the linear condition × length term and the quadratic condition × length-squared term.

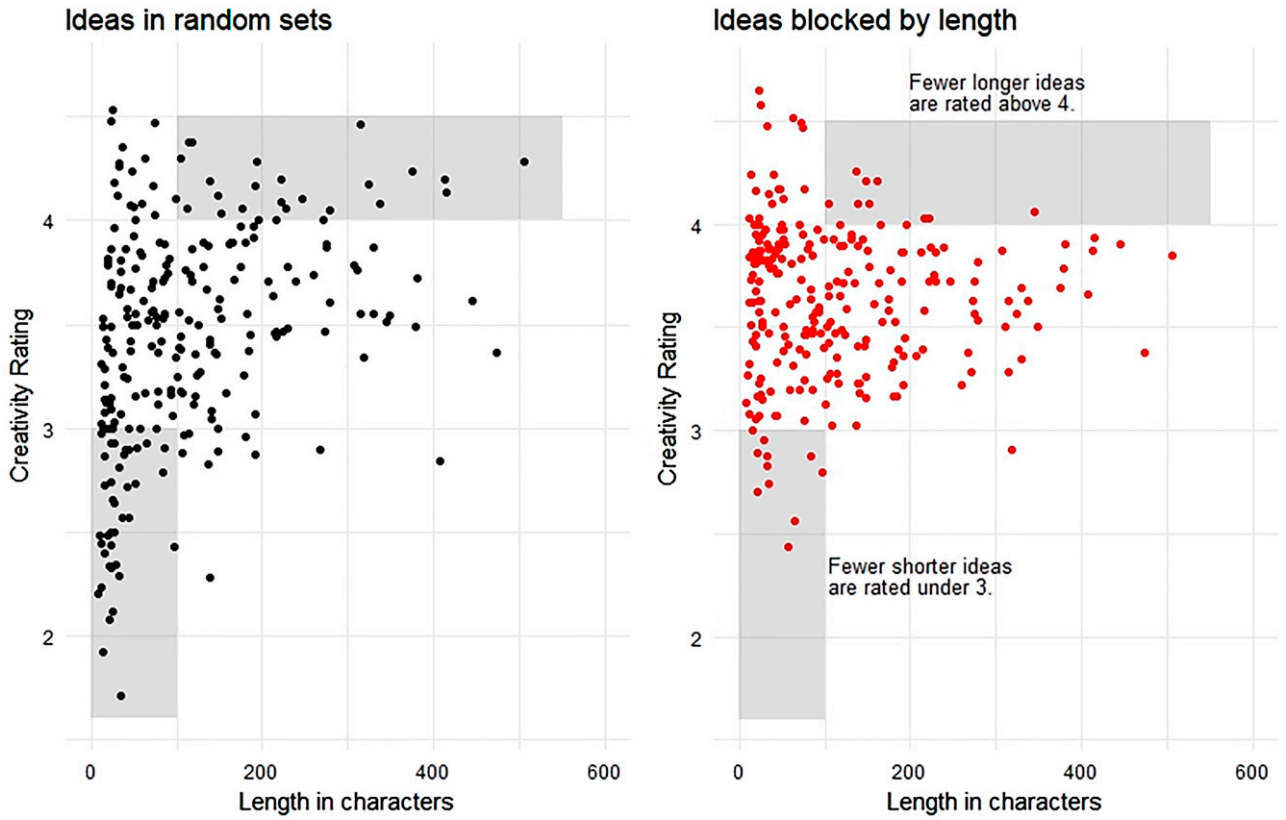
Blocking by length essentially wiped out the relationship between length and creativity. Examining the coefficients in Table 9, we see the wiping out as follows. In the linear model, the slope of the relationship goes from 0.17 in the mixed condition (coefficient on idea length) to 0.17−0.16 in the blocked condition (where −0.16 is the coefficient on condition × idea length). In the quadratic model, the linear and quadratic coefficients go from 0.30 and −0.04, respectively, to (0.30−0.30) and (−0.04 + 0.05).

Discussion

In Study 7, we show that reducing the variation in length in the subset of ideas that people rate reduces (in this study, eliminates) the relationship between length and perceived creativity.

Figure 2 compares the average ratings for the ideas in Study 1 and Study 7. The graph shows the contrast between Study 1, where we showed ideas in random sets, to Study 7, where we showed ideas in blocks of similar length. The shaded boxes in the figure help explain the result. In the lower left shaded boxes: in Study 7, noticeably fewer of the shorter ideas are rated below 3 compared with Study 1. In the upper right shaded boxes: in Study 7, noticeably fewer of the longer ideas are rated above 4. Those two shaded boxes explain why the positive relationship between length and perceived creativity rating disappears in Study 7. The two boxes also provide a clue for the main effect of condition in Table 9. The lower ratings for the

Figure 2. (Color online) Comparison of Ratings of Frosting Ideas in Study 1 and Study 7



Notes. Both panels show length and average creativity rating for 245 of the same 248 frosting ideas in these studies. The three longest ideas are truncated to show detail in the dense part of the graph.

Table 10. Use of Creativity Measures in Marketing Articles About Idea Generation

Journal	Papers with creativity, originality, or novelty as a focal measure	Total papers
J. Consumer Research	1 (25.0%)	4
J. Marketing	5 (71.4%)	7
J. Marketing Research	7 (63.6%)	11
Marketing Science	3 (60.0%)	5
Management Science	2 (15.4%)	13
Total	18 (45.0%)	40

Notes. These results reflect articles that rate product ideas. We searched Web of Science in five journals, from 1999 to 2019, for these terms: "new products," "idea generation," and "ideation." See the full list of papers in Online Appendix C.

longer ideas don't fully offset the higher ratings for the shorter ideas.

This result raises the question of which is the better way to administer the ratings: in sets with lengths drawn from the whole population, or in sets with narrow ranges on lengths? Study 1 and Study 7 alone can't answer that question because in these real ideas, it could be that the longer ideas are more creative ideas. Thus, from those studies alone we can't conclude that eliminating the relationship between length and perceived creativity is an improvement.

Our experiments show that additional length that doesn't change the essential idea does change perceived creativity. Those results tell us that it is appropriate to account for length in some way when comparing the perceived creativity of ideas. One could use a statistical control after ratings are collected: regress creativity on length (and length squared) and use the residual from that regression as the measure of creativity purged of the average length effect. Or one could do what we did in Study 7, present a block of ideas of similar length to each rater.

The approach in Study 7 leverages the underlying mechanism of differentiation. If we stratify ideas by length, raters cannot differentiate them based on length. A statistical approach will work on average, but to the extent that longer ideas are more creative, it will penalize the longer ideas too much (and reward the shorter ideas too much).

General Discussion

In this work, we demonstrate that different ways of expressing the same fundamental idea systematically change the evaluation of that idea. Ideas expressed with longer written descriptions are perceived as more creative than shorter ones. This is true even when the longer version contains no more information than the shorter one, and especially true when the longer version contains concrete even if unsurprising details. Some simple pieces of practical advice for academic researchers and marketing professionals follow from our results.

The first piece of advice is to pay close attention to the variance in description length in a set of raw ideas. If possible, tightly control the range of lengths. When ideas are crowdsourced, there is naturally great variation in style and in length of the descriptions. We can control this variation by enforcing a narrow character range for submitted ideas. Both lower and upper bounds are important. This advice has drawbacks because some ideas just naturally require less description. But with verbal descriptions of raw ideas, all ideas have some detail that can be fleshed out to keep the length uniform. If you can't control the length of the submitted ideas, then control for the length in the evaluations of the ideas. In Study 7, we showed how this change reduced the relationship between length and perceived creativity.

The second piece of advice is to use great care with metrics of creativity or novelty in studies that involve idea generation. For practitioners, base screening in the fuzzy front end on more than just creativity if ideas vary in length. For academic researchers, our results warn against placing too much weight on perceived novelty as a dependent variable, as the perceptions are sensitive to filler in the description or the natural verbosity of some idea writers. In a survey of academic papers on idea generation in top marketing journals, we see that creativity, originality, and novelty are commonly measured dependent variables. Table 10 summarizes those results.

The third piece of advice extends beyond the domain of these studies: if you want someone to think your message is creative or novel, contrast a verbose description with a terse description of the competitor. There are many ways to convey an image of creativity for a brand; the results in this paper suggest that longer descriptions of the product or idea will help.

Surface features matter in many domains. Cosmetic elements of idea descriptions have a real impact on the way the ideas are evaluated. Large innovation enterprises "optimize" ideas and present them in a consistent format before using them in stimuli for research, but this refinement is almost always done after the fuzzy front end, once the set of ideas have been narrowed to a manageable set. We think that is too late in the process, especially if perceived creativity ratings are used to screen ideas into the smaller set.

We hope the insights in this paper make it easier for both academics and marketers to better evaluate which raw ideas hold promise.

Acknowledgments

The authors gratefully acknowledge the helpful input from John Lynch and Nick Reinholtz. They also thank Olivier Toubia, who served as senior editor on this paper and shared the text of ideas from studies in Toubia and Netzer (2017) and to Karl Urlich who shared the text of ideas from Kornish and Ulrich (2011).

Endnote

[1] We built the baseline semantic network from all the ideas across studies 2 and 3 in Toubia and Netzer (2017). For the baseline, we retained stems that appeared in at least five ideas, yielding a baseline semantic network with 389 stems. Across all their studies, TN use a separate source for the baseline, either the text of results from a Google search or the text of a separate sample of the ideas generated for the prompt. Because we selected only 25 ideas out of more than 700, the baseline is reasonably distinct from the set of ideas.

References

Aaker DA (1991) *Managing Brand Equity* (The Free Press, New York).

Acar S, Burnett C, Cabra JF (2017) Ingredients of creativity: Originality and more. *Creativity Res. J.* 29(2):133–144.

Alba JW, Broniarczyk SM, Shimp TA, Urbany JE (1994) The influence of prior beliefs, frequency cues, and magnitude cues on consumers' perceptions of comparative price data. *J. Consumer Res.* 21(2):219–235.

Alexander DL, Lynch JG Jr, Wang Q (2008) As time goes by: Do cold feet follow warm intentions for really new vs. incrementally new products? *J. Marketing Res.* 45(3):307–319.

Åstebro T, Elhedhli S (2006) The effectiveness of simple decision heuristics: Forecasting commercial success for early-stage ventures. *Management Sci.* 52(3):395–409.

Berg JM (2019) When silver is gold: Forecasting the potential creativity of initial ideas. *Organ. Behav. Human Decision Processes* 154:96–117.

Berlyne DE (1960) *Conflict, Arousal, and Curiosity* (McGraw-Hill, New York).

Berlyne DE (1970) Novelty, complexity, and hedonic value. *Perception Psychophysics* 8(5):279–286.

Broniarczyk SM, Griffin JG (2014) Decision difficulty in the age of consumer empowerment. *J. Consumer Psych.* 24(4):608–625.

Brown CL, Carpenter GS (2000) Why is the trivial important? A reasons-based account for the effects of trivial attributes on choice. *J. Consumer Res.* 26(4):372–385.

Carpenter GS, Glazer R, Nakamoto K (1994) Meaningful brands from meaningless differentiation: The dependence on irrelevant attributes. *J. Marketing Res.* 31(3):339–350.

Chaiken S, Maheswaran D (1994) Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *J. Personality Soc. Psych.* 66(3):460–473.

Diedrich J, Benedek M, Jauk E, Neubauer AC (2015) Are creative ideas novel and useful? *Psych. Aesthetics Creativity Arts* 9(1): 35–40.

Eliashberg J, Hui SK, Zhang ZJ (2007) From story line to box office: A new approach for green-lighting movie scripts. *Management Sci.* 53(6):881–893.

Girotra K, Terwiesch C, Ulrich KT (2010) Idea generation and the quality of the best idea. *Management Sci.* 56(4):591–605.

Goldenberg J, Lehmann DR, Mazursky D (2001) The idea itself and the circumstances of its emergence as predictors of new product success. *Management Sci.* 47(1):69–84.

Goldenberg J, Mazursky D, Solomon S (1999) Toward identifying the inventive templates of new products: A channeled ideation approach. *J. Marketing Res.* 36(2):200–210.

Graf LK, Mayer S, Landwehr JR (2018) Measuring processing fluency: One vs. five items. *J. Consumer Psych.* 28(3):393–411.

Grice HP (1975) Logic and conversation. Cole P, Morgan JL, eds. *Speech Acts* (Brill, Leiden, Netherlands), 41–58.

Hoeffler S (2003) Measuring preferences for really new products. *J. Marketing Res.* 40(4):406–420.

Kauffman SA (1984) Emergent properties in random complex automata. *Physica D* 10(1–2):145–156.

Kornish LJ, Ulrich KT (2011) Opportunity spaces in innovation: Empirical analysis of large samples of ideas. *Management Sci.* 57(1):107–128.

Kornish LJ, Ulrich KT (2014) The importance of the raw idea in innovation: Testing the sow's ear hypothesis. *J. Marketing Res.* 51(1):14–26.

Kudrowitz BM, Wallace D (2013) Assessing the quality of ideas from prolific, early-stage product ideation. *J. Engrg. Design* 24(2):120–139.

Levinthal DA (1997) Adaptation on rugged landscapes. *Management Sci.* 43(7):934–950.

Lurie NH (2004) Decision making in information-rich environments: The role of information structure. *J. Consumer Res.* 30(4): 473–486.

Morewedge CK, Kahneman D (2010) Associative processes in intuitive judgment. *Trends Cognitive Sci.* 14(10):435–440.

Mugge R, Dahl DW (2013) Seeking the ideal level of design newness: Consumer response to radical and incremental product design. *J. Product Innovation Management* 30(S1):34–47.

Mukherjee A, Hoyer WD (2001) The effect of novel attributes on product evaluation. *J. Consumer Res.* 28(3):462–472.

Nunes JC, Ordanini A, Valsesia F (2015) The power of repetition: Repetitive lyrics in a song increase processing fluency and drive market success. *J. Consumer Psych.* 25(2):187–199.

Page SE (2015) What sociologists should know about complexity. *Annual Rev. Sociol.* 14(41):21–41.

Petty RE, Cacioppo JT (1984) The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *J. Personality Soc. Psych.* 46(1):69–81.

Poetz MK, Schreier M (2012) The value of crowdsourcing: Can users really compete with professionals in generating new product ideas? *J. Product Innovation Management* 29(2):245–256.

Preacher KJ, Hayes AF (2004) SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav. Res. Methods Instruments Comput.* 36(4):717–731.

Reber R, Winkielman P, Schwarz N (1998) Effects of perceptual fluency on affective judgments. *Psych. Sci.* 9(1):45–48.

Rogers EM (1995) *The Diffusion of Innovations*, 3rd ed. (The Free Press, New York).

Rosenzweig S, Tellis GJ, Mazursky D (2015) Where does innovation start: With customers, users, or inventors? Marketing Science Institute Working Paper Series 2015 Report No. 15-108, Marketing Science Institute, Cambridge, MA.

Runco MA, Jaeger GJ (2012) The standard definition of creativity. *Creativity Res. J.* 24(1):92–96.

Schwarz N (2004) Metacognitive experiences in consumer judgment and decision making. *J. Consumer Psych.* 14(4):332–348.

Schwarz N (2010) Meaning in context: Metacognitive experiences. Mesquita B, Barrett LF, Smith ER, eds. *The Mind in Context* (Guilford, New York), 105–125.

Scott EL, Shu P, Lubynsky RM (2020) Entrepreneurial uncertainty and expert evaluation: An empirical analysis. *Management Sci.* 66(3):1278–1299.

Sela A, Berger J (2012) How attribute quantity influences option choice. *J. Marketing Res.* 49(6):942–953.

Shafir E, Simonson I, Tversky A (1993) Reason-based choice. *Cognition* 49(1–2):11–36.

Stang DJ (1977) On the relationship between novelty and complexity. *J. Psych.* 95(2):317–323.

Thompson DV, Hamilton RW, Rust RT (2005) Feature fatigue: When product capabilities become too much of a good thing. *J. Marketing Res.* 42(4):431–442.

Toubia O, Netzer O (2017) Idea generation, creativity, and prototypicality. *Marketing Sci.* 36(1):1–20.

Tversky A (1977) Features of similarity. *Psych. Rev.* 84(4):327–352.

Zhao X, Lynch JG Jr, Chen Q (2010) Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *J. Consumer Res.* 37(2):197–206.