Evidence of Muddy Knowledge in Reaching for the Stars: Creating Novel Endings for Event Sequences

Rebecca Grimes-Maguire (rebecca.grimes@ucd.ie)
Mark T. Keane (mark.keane@ucd.ie)
Department of Computer Science, University College Dublin
Belfield, Dublin 4, Ireland

Abstract

This experiment examines people's ability to invent creative outcomes to simple event sequences. We report a study where participants are given everyday event descriptions and asked to describe either a predictable outcome (Predictable group) or a creative outcome (Creative group). Following the Creative Cognition approach (Finke, Ward & Smith, 1992), we expected that though those instructed to be creative might generate novel and interesting outcomes, they would also be bound by their knowledge of the outcomes that typically occur. The results support this prediction, in that while the Creative group manifested more inventive variability in their outcomes relative to the Predictable group, their proposed outcomes still overlapped in part with those of the Predictable group. These results show that although creativity may take people beyond their knowledge, they can never fully break free from that knowledge.

Introduction

Creativity is like reaching for the stars with your feet firmly stuck in the mud of everyday life. While the creative act takes us beyond what an individual or a society has thought before, it seems to be inextricably constrained by what is known already (see e.g., Boden, 1995; Finke, Ward & Smith, 1992; Perkins, 1981; Sternberg, 1999). In science, new theories come from reactions to old paradigms, but still work from the same methodologies and findings of previous decades. In the arts, similar reactions to the conventions of a previous age occur, though often themes and materials remain the same. In this paper, we examine this interplay between creativity and the constraints placed on it by prior knowledge by studying people's generation of novel outcomes to conventional event sequences. We often need to imagine unconventional or novel outcomes to typical happenings (e.g., in launching a new product or assessing the impact of new technologies). Yet, we know of little work which examines people's creativity in such situations.

The idea that creativity is often constrained by prior knowledge has been strongly and convincingly argued for in Finke et al.'s (1992) Creative Cognition Approach and their 'Geneplore model'. This approach has been supported by several appropriate empirical demonstrations. For example, Ward (1994) asked participants to imagine and then draw creatures that live on a distant planet very unlike earth. This simple creative task revealed that converse to the instructions, almost all the participants produced animals very like the ones living on this planet, in that they exhibited features such as bilateral symmetry, external organs (e.g. eyes, ears) and appendages (e.g. legs, tails). Ward concluded that the participants were constrained by their experience of real world animals and could not deviate easily from such prototypes. This type of influence from background knowledge has been subsequently demonstrated in studies on the generation of novel product names (Rubin, Stoltzfus & Wall, 1991), ideas for traffic improvement (Marsh, Landau & Hicks, 1997), and even the generation of non-words (Marsh, Ward & Landau, 1999). Haught & Johnson-Laird (2003) have reported similar findings in a task where people were asked to come up with a creative sentence incorporating two or three specific nouns. They found that people were quite restricted in their output; for example, the words 'lion' 'strawberry' and 'harp' tended to result in similar sentences, such as "The lion was playing the harp while eating the strawberry". In Ward's (1994) view, these "structured imagination" effects occur because, when faced with a problem whose solution requires creativity, people tend to take the path of least resistance by retrieving domain specific information or an existing solution (whether this is an experimenter provided example or self-generated from previous knowledge) and then attempt to modify the old construct in some novel way.

In the present study, we look at the constraint placed by background knowledge in a task that deals with novel sentence generation involving script-like scenarios (Bower, Black & Turner, 1979; Schank & Abelson, 1977). In our task, people are presented with typical event sequences that have incomplete, but predictable, endings (see Appendix). These scenarios either involve conventional events that proceed uninterrupted (e.g. "Matthew had wanted to quit his job for months. One day he walked into his boss's office…"), or events that are interrupted by some surprising event or state (e.g. "The cup of coffee was balanced on the arm of the chair. Suddenly, Richard sneezed….). In both cases, the main manipulation was to ask people to come up with a creative outcome to the sequence. In the remainder of this paper, we detail this study and sketch the properties of a computational model that might capture the effects found.

Proposing Creative Outcomes to Events

Following the Creative Cognition Approach, we expected that our creative-ending generation task would manifest the constraining influence of prior, background knowledge in our participants. In the experiment, there were two main groups, the Predictable and Creative groups. Both groups were given the same set of event sequence materials

(divided into Unfolding and Surprise scenarios). However, the Predictable group was asked to "think of a typical ending to the scenario…", whereas the Creative group was asked to "think of a creative ending to the scenario…". Thus the design was a 2 x 2 one, with Group being a between-participants variable (Predictable or Creative) and Scenario being a within-participants variable (Unfolding or Surprise).

The main prediction was that the Creative group would generate many of the same outcomes as the Predictable group, as they would be constrained by their background knowledge of the typical endings of these events. However, we thought that something additional would also be included in these endings, giving them an added novel twist. So, in the specific measures we used, we expected more elaborate endings in the Creative group (i.e., more propositions generated), but we also expected that some of these propositions would overlap with those produced by the Predictable group (i.e., propositions reflecting a common ending). To put it another way, the Predictable group's endings should strongly overlap with those generated by the Creative group.

We had no apriori grounds for expecting a difference between the Unfolding and Surprise scenarios, though they do appear to be distinct categories. In the Unfolding scenarios the sequence of actions proceeds unchecked in a predictable way. In the Surprise scenarios one state or sequence of actions is cut across by another sequence of actions. Interestingly though, in the Surprise scenarios the interrupting sequence is also predictable, it's a "typical surprise" (e.g., a poorly balanced object being knocked).

Experiment

Method

Participants Thirty native English-speaking undergraduate psychology students from University College Dublin volunteered for this experiment.

Materials Twenty-four scenarios involving typical everyday event sequences (see Appendix). All scenarios had two sentences and required a third to complete the sequence. The 24 materials consisted of three types of sequences: 8 Unfolding items, 8 Surprise items and 8 filler items. The Unfolding items described two events/states in a typical sequence with a predictable outcome (e.g. "Cathy saw the cake in the window. She hadn't had lunch that day…"). The Surprise items described one event/state that was interrupted by another event/state leading to a predictable outcome (e.g., "The little boy played by the edge of the pond. Suddenly he slipped on some moss…").

Design In the 2 x 2 (Group x Scenario) design, participants were randomly assigned to one of two between-participant groups, Predictable (N=15) and Creative (N=15). All participants received the same 24 scenarios, which were presented in a different random order to each participant.

Procedure Each participant was given a booklet containing all the materials, the first page of which included instructions. The items were presented so that only one scenario appeared on each page (one sentence per line with a prompt stating 'your ending:' in the space below each scenario). Participants in the Predictable group were given the instructions to "think of a typical ending to the scenario", whilst those in the Creative group were asked to "think of a creative ending" to be written as a concluding sentence. In the Creative group participants were also asked to describe a "creative turn of events, not just the use of creative language" so as to avoid a misinterpretation of the instructions. An earlier pilot showed that without this instruction, some people just produced purple-prose versions of typical endings rather than truly novel endings.

Scoring Participants' responses in completing the presented sentences were firstly rated for level of *creativity*. Then the responses were analysed into propositions. As a further measure of creativity, we wished to examine the *diversity* and *richness* of responses made, but we also examined the *commonalities* between responses to determine if there was any overlap across the different conditions.

To measure *creativity*, following Haught & Johnson-Laird's (2003) procedure, two judges independently rated each sentence (blind to condition) on a 7-point scale, with a score of 1 denoting a highly uncreative sentence and a score of 7 denoting an extremely creative sentence.

To measure *diversity*, for each item we categorised the distinct propositions used in people's endings. So for example, for the "Cathy looked at the cake in the shop window. She hadn't had lunch that day…" item, there were three distinct classes of responses given as endings:

(1) Cathy gets the cake.
(2) Cathy decides not to get the cake for some reason (e.g., diet).
(3) Cathy was hungry.

To measure the *richness* of the responses, we scored the endings produced for their word length and the number of different events mentioned in them. This measure was used because, even though people were asked to provide just one sentence, in many cases multiple events/states were included in the responses. So for example, in the cake-seeing scenario, the response "But she knew she was on a diet so decided to wait until she got back to her office, and then ate something less fattening", was classed as having *three* events/states:

(1) Attribute of being on a diet (a state).
(2) Cathy went back to office (event 1).
(3) She ate something less fattening (event 2).

To measure *commonality*, we noted the most common response, i.e. the frequency of occurrence of a given response across a given group. In the cake-seeing scenario the most common event was "Cathy gets the cake" which received 11 counts in the Predictable group and 9 in the Creative group.

For each of these measures, two raters independently scored the materials. The inter-rater reliability was uniformly high on each; for example, in the diversity measure a random sample of ratings showed 94.99% inter-rater reliability in categorising the different responses.

Table 1: Sample responses from two scenarios

"Katie searched everywhere for her little kitten. Then she heard a miaow from the bin"

Predictable	*N*	*Creative*	*N*
Katie opened bin	11	Katie opened bin	4
An explanation that kitten was in bin	3	An explanation that kitten was in bin	4
Kitten walked out of bin	1	Katie found whole cluster of kittens	1
		Katie found wrong kitten	1
		Kitten was being carried away	1
		Katie couldn't understand how kitten was in bin	1
		Bin collector came	1
		Katie was relieved	1
		Katie was disappointed	1

"The cup of coffee was balanced on the arm of the chair. Suddenly Richard sneezed"

Predictable	*N*	*Creative*	*N*
Cup of coffee fell	13	Cup of coffee fell	7
Richard saved cup from falling	1	Richard saved cup from falling	1
Richard's snot went into coffee	1	Richard's snot went into coffee	3
		Chair propelled backwards	1
		Friend got shock and dropped her cup of hot chocolate	1
		Spaceship flew out of nose	1
		A gust of wind went through the window	1

Results & Discussion

To summarise, analysis of the results showed that, though the Creative group produced more creative, diverse and richer responses than the Predictable group, they also could not avoid the commonly occurring events that were invited by the scenario. These results demonstrate that in generating novel outcomes, people are restricted by their background knowledge. Table 1 illustrates samples of the responses made by participants in two scenarios in the experiment.

Creativity of Responses All of the responses were rated blind-to-condition by two judges independently on a 7-point scale. The judges' ratings were reliably correlated (Pearson's $r = 0.748$, $p < 0.01$). A 2 x 2 ANOVA on these ratings for the Group (between-participants) and Scenario (within-participants) variables revealed a main effect of Group, Materials and a reliable interaction, $F(1,478) = 4.65$, $p < 0.05$, $MSe = 6.01$. As expected, responses from the Creative group ($M = 3.523$) were rated as being more creative than those of the Predictable group ($M = 2.245$), $F(1,478) = 188.82$, $p < 0.01$, $MSe = 397.84$. It was also found that the Unfolding materials were rated as more creative ($M = 2.96$) than the Surprise materials ($M = 2.80$), $F(1, 478) = 4.897$, $p < 0.05$, $MSe = 6.338$. This finding suggests that such unfolding events promote greater creative products. This was an unexpected result. It may indicate that it is harder to break the inevitability of the outcome to a surprise scenario because its outcome is much more highly determined. However, we should exercise some care in making sweeping conclusions about this difference, as it is not reflected in any of the other measures. The reliable interaction showed that the most creative condition was the Creative-Unfolding one ($M = 3.68$), followed by the Creative-Surprise ($M = 3.36$), Predictable-Unfolding ($M = 2.24$) and Predictable-Surprise ($M = 2.23$) conditions respectively.

Diversity of Responses Overall, one would expect a greater diversity of responses in the Creative group versus the Predictable group. This is exactly what we found (see Figure 1). A 2 x 2 ANOVA on the diversity scores revealed a reliable main effect of Group, with the Creative group generating more classes of responses ($M = 6.37$) than the Predictable group ($M = 3.5$), $F(1, 14) = 34.6$, $p < 0.01$, $MSe = 66.13$. There was no reliable effect of Scenario and no reliable interaction. An indication of the greater diversity in Creative responses can be seen in Table 1 where both scenarios show more diversity in the Creative condition.

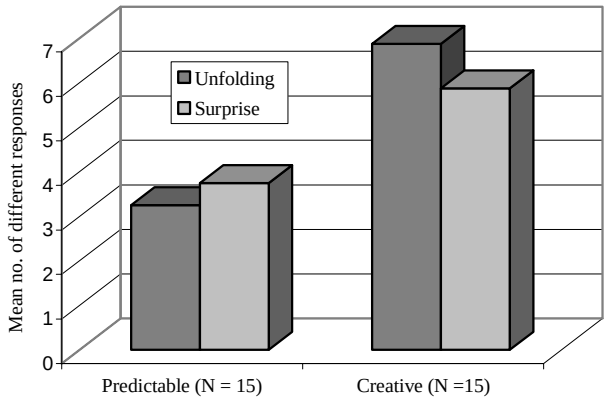

Figure 1: Diversity - the mean number of different responses generated for each condition

Richness of Responses Another index of creativity is the elaborateness or richness of the endings generated. In general, one would expect a greater richness in the responses made by the Creative group than by the Predictable group. Like Haught & Johnson-Laird (2003) we tapped this dimension by examining the average sentence length of people's endings. In addition to this we calculated the mean number of different events in each response.

A 2 x 2 ANOVA again revealed a main effect of Group but no other reliable effects. The Creative group was more likely to provide longer responses ($M = 12.5$ words) than the Predictable group ($M = 9.24$ words), $F(1, 233) = 35.071$, $p < 0.01$, $MSe = 1220.29$. An example of a Predictable response for the first scenario in Table 1 was "She reached in and pulled the kitten out" (word count = 8), a creative response for the same scenario was "She pulled a white kitten from the bin, her kitten was black so she put the white kitten back and carried on looking" (word count = 23).

A 2 x 2 ANOVA on the mean number of different events in the ending showed a comparable pattern; a reliable main effect of Group, but no other effects. The Creative group was more likely to include additional events per ending (M = 1.8) than the Predictable group (M = 1.46), $F(1,233)$ = 19.583, $p < 0.01$, MSe =13.86. Using the example above, the predictable response was classed as having one event and the creative was classed as having three events.

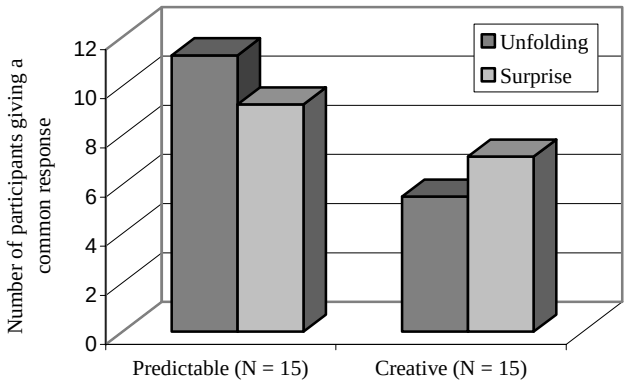

Figure 2: Commonality – count of most common response for each scenario across conditions

Commonality of Responses The above measures show the generativity of the Creative group at work relative to the Predictable one, but they do not reveal the constraints we expected from background knowledge based on Ward's (1994) proposals. If this constraint is in evidence we should see that, in spite of the clear differences in the creativity of responses, there should be certain commonalities between the Creative and Predictable groups too. Specifically, we should see many of the Creative group using the same, inevitable events as some part of their endings. As Figure 2 illustrates, this is exactly what we found. In an analysis of the most commonly produced response, we observed that while those in the Predictable condition (M = 10.25) were more likely to produce the common event, those in the Creative condition (M = 6.313) also produced this same event to a high degree, $F(1, 14) = 31.042$, $p < 0.01$, MSe = 124.03). Again an example of this can be seen in the second scenario in Table 1: the most common response for all participants is that the "cup of coffee fell", which received high counts in both conditions.

General Discussion

The aim of this study was to explore how creative instructions would influence an individual's completion of a common event sequence. The Creative Cognition approach argues that background knowledge can play a constraining role in the creative process, a proposal that is confirmed by the present results. More specifically this experiment has shown that the expectations we have about certain events in the world have a profound influence on our thought processes. We found that although responses of the Creative group were more creative, rich and diverse than those in the Predictable group, certain elements of the endings provided by both groups overlapped considerably. Thus, it appears that whilst creativity in essence involves some degree of variability and unpredictability, it is firmly rooted in our background knowledge of events. In the remainder of this section, we discuss the relationship of these results to the literature on comprehension, and how they might be modeled computationally.

Consistency With Theories of Comprehension

Graesser Singer & Trabasso (1994) stress that knowledge of goals, actions and events are deeply embedded in our perceptual and social experience. As we interact with the environment, we have a strong tendency to interpret event sequences as causal seqeunces, and a similar process occurs in comprehension by means of *inferences* (Kintsch, 1998; Zwaan, 1999). For example, Duffy (1986) observed that when reading, we continually form expectations about future events, so as to develop a causal chain of narrative. The 'Situation Model' of comprehension holds that we construct a detailed mental representation of people, objects, locations, events and actions described in a text (e.g. Zwaan, 1999). Consequently, when reading the scenarios of the present experiment, it was difficult for participants to avoid making rapid, almost automatic inferences about the mental states of the characters and/or the events that would typically occur. In the scenario where Cathy sees the cake for example, it can easily be inferred that Cathy is hungry and that she would like to eat the cake. It could be hypothesised that in this task, the participants naturally link the two sentences together, and in order to provide a coherent ending, they must fit their response with the depicted events so that it is easily understandable and 'makes sense' when read. It is this fundamental knowledge constraint that often overrides instructions to be creative.

Possible Computational Models

Connell & Keane (2002, 2003, in press) have developed a computational model of plausibility judgements for event sequences that is consistent with the above general theory of comprehension. At present, this model takes some event description and finds alternative possible inferential paths to link the events described, this elaborated representation then being scored to assess the plausibility of the description. As such, this Plausibility Analysis Model (PAM) is one possible candidate model that could be extended to deal with the present findings. Such an extension would have to rely on two significant changes: (i) the generation of further possible events to a given sequence based on background knowledge and then (ii) the application of some set of selection heuristics to rank order these possible outcomes for their novelty. As a first pass, such rules could just favour less likely outcomes; that is, outcomes that could possibly occur but that are not strongly supported by prior experience. Obviously such an extension would invite new predictions about the relationship between plausibility and creativity too.

Concluding Comments

The present paper reports a novel study of people's ability to generate creative endings to sentences describing commonplace event sequences. This work connects several areas that have previously been quite separate; namely creativity, sentence comprehension and plausibility. The convergence of these three areas presents a real opportunity for understanding this type of creativity in a new and computationally well-specified way. In short, we should be able to characterise the mud of everyday knowledge, exactly how it glues us to the ground and, yet, the exact nature of the way in which we reach for the stars.

Acknowledgments

This research was funded by a grant to the first author from the Irish Research Council for Science, Engineering and Technology and supported by Science Foundation Ireland under Grant No.03/IN.3/I361 to the second author. Thanks to Phil Maguire and three anonymous reviewers who provided some helpful comments.

References

Boden, M. A. (1995). Creativity and unpredictability. *Stanford Education and Humanities Review, 4* (2).

Bower, G. H., Black, J. B., & Turner, T. (1979). Scripts in memory for text. *Cognitive Psychology, 11*, 177-220.

Connell, L., & Keane, M.T. (2002). The roots of plausibility: The role of coherence and distributional knowledge in plausibility judgments. *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Connell, L., & Keane, M.T. (2003). PAM: A cognitive model of plausibility. *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Connell, L., & Keane, M.T. (in press). What plausibly affects plausibility: Concept coherence and distributional word coherence as factors influencing plausibility judgments. To appear in *Memory and Cognition.*

Duffy, S. A. (1986). Role of expectations in sentence integration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 208-219.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative Cognition: Theory, Research, and Applications.* Cambridge. MA: Bradford-MIT Press.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395.

Haught, C., & Johnson-Laird, P. N. (2003). Creativity and constraints: The creation of novel sentences. *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition.* Cambridge: Cambridge University Press.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 886-897.

Marsh, R. L., Ward, T. B., & Landau, J. L. (1999). The inadvertent use of prior knowledge in a generative cognitive task. *Memory & Cognition, 27*, 94-105.

Perkins, D.N. (1981). *The Mind's Best Work.* Cambridge: Harvard University Press.

Rubin, D. C., Stoltzfus, E. R., & Wall, K L. (1991). The abstraction of form in semantic categories. *Memory and Cognition, 19*, 1-7.

Schank, R. C., & R. P. Abelson. (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (Ed.). (1999). *Handbook of Creativity*. Cambridge, MA: Cambridge University Press.

Ward, T. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology, 27*, 1-40.

Zwaan, R. (1999). Situation models: The mental leap into imagined worlds. *Current Directions in Psychological Science, 8*, 15-18.

Appendix: Materials Used in the Study

Unfolding Scenarios

1 Cathy looked at the cake in the shop window. She hadn't had lunch that day.
2 The dog saw the bone in kitchen bin. He wagged his tail in anticipation.
3 Thomas the cat felt bored. He noticed the dangling tablecloth.
4 Katie searched everywhere for her little kitten. Then she heard a miaow from the bin.
5 James wanted to read the paper. He stopped at the shop on his way home from work.
6 Robert hated his old car. He decided to call the bank.
7 Matthew had wanted to quit his job for months. One day he walked into his boss's office.
8 Jim felt very cold. He got some coal and firelighters.

Surprise Scenarios:

1 John and Pat were kicking a football on the street. A speedy car sharply turned the corner.
2 Michael's shopping bags were bursting with groceries. He felt one of the handles begin to break.
3 The cup of coffee was balanced on the arm of the chair. Suddenly, Richard sneezed.
4 The yacht sailed on as the crew slept. A rocky reef lay directly ahead.
5 Paul was crossing a busy road. Unexpectedly his mobile phone rang.
6 Peter and Sally ate lunch in the small restaurant. They didn't realise that the meat wasn't properly defrosted.
7 The little boy played at the edge of the pond. Suddenly, he slipped on some moss.
8 The sheep were grazing in the field. Suddenly, a wolf approached the flock.