

Interactive Skill in Scrabble

Paul P. Maglio (pmaglio@almaden.ibm.com)
IBM Almaden Research Center
650 Harry Rd., NWED-B2
San Jose, CA 95120 USA

Teenie Matlock (tmatlock@cats.ucsc.edu)
Psychology Department
University of California, Santa Cruz
Santa Cruz, CA 95064 USA

Dorth Raphaely (sunnyboy@cats.ucsc.edu)
Psychology Department
University of California, Santa Cruz
Santa Cruz, CA 95064 USA

Brian Chernicky (brian@tapestry.net)
Psychology Department
University of California, Santa Cruz
Santa Cruz, CA 95064 USA

David Kirsh (kirsh@ucsd.edu)
Cognitive Science Department
University of California, San Diego
La Jolla, CA 92093 USA

Abstract

An experiment was performed to test the hypothesis that people sometimes take physical actions to make themselves more effective problem solvers. The task was to generate all possible words that could be formed from seven Scrabble letters. In one condition, participants could use their hands to manipulate the letters, and in another condition, they could not. Results show that more words were generated with physical manipulation than without. However, an interaction was obtained between the physical manipulation conditions and the specific letter sets chosen, indicating that physical manipulation helps more for generating words in some circumstances than in others. Overall, our findings can be explained in terms of an interactive search process in which external, physical activity effectively complements internal, cognitive activity. Within this framework, the interaction can be explained in terms of the relative difficulty of generating words from the letters given in the different sets.

Introduction

People often adapt their physical environments to take better advantage of cognitive or perceptual skills (Clark, 1997; Hutchins, 1995a; Kirsh & Maglio, 1994; Kirsh, 1996). For instance, Tetris players take actions to set up their external environment to facilitate perceptual processing (Kirsh & Maglio, 1994; Maglio & Kirsh, 1996); gin rummy players physically organize the cards they have been dealt so as to be able simply to read off what is in a hand (Kirsh, 1995);

and airline pilots place external markers on their controls to help keep track of appropriate speed and flap settings (Hutchins, 1995b). In each of these cases, people take action to set up their external environments so that their mental jobs are easier, faster, or less error-prone.

The key to such processes of *interactive skill* is that the benefits of adapting the physical environment outweigh the costs of taking the physical actions. In the case of Tetris, for example, the cost of rotating a falling Tetris piece too many times is small (because over-rotation can quickly be corrected) compared to the benefit of relying on the visual system to determine whether the piece fits in its visible orientation. Thus, the hypothesis is that it is more efficient to rotate the falling piece in the external environment than to imagine how it would look in a different orientation.

Most people are familiar with the board game Scrabble¹, in which players form words by arranging tiles with letters printed on them. When trying to come up with words in this game, people can either mentally rearrange the letters or physically rearrange the letters. Based on the idea that people routinely set up their environments to make their cognitive jobs easier, it is reasonable to suppose that it is easier to form words by physically moving the tiles than by simply imagining their rearrangement (Kirsh, 1995). But is this really true? If so, is it always the case? Our first objective

¹ "Scrabble" is a registered trademark of Hasbro, Inc.

is to test the conjecture that the physical action of moving Scrabble tiles facilitates the discovery of anagram solutions. More precisely, given a sequence of seven letters, such as “RDLOSNA”, and the task of calling out all legal words containing at least two letters in five minutes, will more words be formed in the condition in which the tiles can be moved than in the condition in which the tiles cannot be moved? As we will show, the answer is yes.

Overall, we found that participants generated more words when they were allowed to manipulate the Scrabble tiles than when they were not. However, we also found that this was not always the case. In particular, physically arranging the letters led to more words for only one of the two sets of letters tested. Though we attempted to control for productivity of the letter sets through a norming task, it turned out that the words found in one of the sets are far more frequent in English than the words found in the other. Use of hands facilitated word generation only for the less frequent set. It is reasonable to suppose that less frequent words are harder to generate and so would more likely benefit from any external help.

From a theoretical perspective, it makes sense that physically arranging letters simplifies the task of forming words, as it ought to be easier to see words by looking than to see words by mentally swapping letters around. But there are many possible ways a person might generate words when given a set of letters. Our modeling objective is to discover an underlying process model to explain our finding that people form more words when they can move the Scrabble tiles than they form when they cannot.

Clearly, people engage in a search process of some sort, but we are not sure of the state-space that they are searching, the operators that yield neighboring states, or the subjective metric that is used to judge states. The most obvious state-space search is one in which the states represent letter strings and the transitions between the states represent the operations of adding, deleting, and swapping letters. In such a state-space, for instance, “chat” might be found from “hat” by adding a “c” to the front of the string. An alternative state-space might contain operators that can move from one state to another along a semantic or associative dimension; for instance. In this case, “cat” might be found from “hat” because “cat” is associated with “hat” in the familiar title, “The Cat in the Hat”. In any event, the model must account for how the external actions of manipulating letters can have the cognitive effect of improving performance, especially for the set of less frequent words.

In what follows, we first sketch a model of interactive skill in Scrabble, and then describe our empirical study.

Model of Interactive Scrabble Skill

One way to think about the process of generating Scrabble words is in terms of the metaphor of energy landscapes. If we regard the set of legal words created from a letter sequence seven letters long to be a set of attractors distributed in a state-space consisting of letter sequences between two and seven letters long, then we can interpret the search for words to be some sort of stochastic hill-climbing process. The energy metric for this landscape might be determined,

for example, by frequency of bigrams, trigrams, and words, as well as the probability that a bigram such as **br** will be continued with an **e**, or continued with an **o**, and so forth. Given such a landscape, we can then attempt to explain both the timing and sequence of the anagram solutions that participants provide by suggesting that they engage in a particular type of search.

In fact, if we assume some sort of stochastic search, the reason hands help would be the same regardless of the details of the model. Specifically, physical manipulation allows one to instantly jump to new parts of the state space and to begin searching there. In a sense, the mental search for words is hampered by the data-driven nature of looking at the tiles. Consider the analogy of a rubber band: People can generate diverse letter combinations in their heads, but when they re-examine the tiles, they are drawn back to the original arrangement, like a rubber band springs back to its original shape. Thus, it is hard to continue searching from positions in the search space that are distant from the visible arrangement of the tiles. If words are not too difficult to find, they can be discovered quickly without having to look again at the tiles. If words are difficult to find, however, it may be helpful to be reminded of the letters by looking at them.

Nevertheless, to make good on this sort of search-based model, we must specify in detail not only the operators that define the state-space, but also the energy landscape of the space. We are just now beginning to explore such a model, so we can only sketch it in the broadest strokes. For instance, we do not yet know how to combine information about the frequency of words and their parts into a single number that gives the “closeness” – the energy level – of a state. Nor can we make precise the notion of locality; that is, how to decide when two states are neighbors (regardless of how distant they are in energy terms). Then there is the question of defining the state-space itself; for instance, should it be defined over letter sequences, using the operators of “add”, “delete”, “substitute”, “rearrange” as primitives? And there is the problem of words that cannot be formed from the allotted letters but that are attractive because they have a similar sound or meaning to words that can be formed from the allotted letters.

To start, suppose we choose the following operators,

ear ⇒ **bear** **ore** ⇒ **ogre** – *arbitrary add*
ago ⇒ **age** **boor** ⇒ **boer** ⇒ **boar** – *arbitrary substitute*
bore ⇒ **ore** **brag** ⇒ **bag** – *arbitrary delete*
ogre ⇒ **gore** **bear** ⇒ **bare** – *arbitrary rearrange*

Taken together, these seem too powerful. After all, there is a cost to mental search, and to performing computations in working memory. Perhaps it would be appropriate to restrict “add”, “delete”, and “substitute” operators so that they only apply to the beginning or end of a word (i.e., no center embedding in a single step). To achieve arbitrary letter strings, then, these operators would have to be combined with “rearrange”. Perhaps we should include “reverse” as a special type of rearrangement that permits more global changes in a single move, such as

gob ⇒ **bog** **garb** ⇒ **brag** – reversal, special rearrange

Of course, people are not dealing with abstract and arbitrary strings when playing Scrabble. Strings form syllables, syllables form words, and so on. In arranging letter strings, people must contend with word-formation constraints of English, including permissible consonant clusters, consonant-vowel sequences, or vowel-vowel sequences. For example, **rd** and **dr** are both permissible consonant clusters in English; however, **rd** is not allowed in word-initial position, and **dr** is not allowed in word-final position. Such orthographic and phonetic constraints are far too numerous and complex to list here.

A similar issue concerns the units the operators act on. These units might be restricted to individual letters, so that only one letter can be altered in a single action, as in **ogre** ⇒ **gore**, or they may be able to operate on entire letter sequences (**bear** ⇒ **bare**). Perhaps the operators ought to be restricted to bigrams at the beginning or ending of a string, as in

bar ⇒ **barge** – append bigram
rage ⇒ **gear** – rearrange bigram

In this first pass analysis, however, such questions about operators are of less consequence than the metric that determines how close (easy to reach) neighbors are. Accordingly, even if we are wrong in supposing that **garb** is an immediate neighbor of **brag** in state-space, the key factor determining whether **brag** is generated soon after **garb**, is principally a function of their closeness in terms of energy – that is, the relative difficulty of making the transition from one state to another. In this way, our modeling approach is similar to Hofstadter’s (1983) Jumbo program, which used relatively little knowledge of English to solve anagram puzzles through a stochastic search of the space of letter and syllable clusters. In such a stochastic system, the primary means for assuring that enough of the space is searched (i.e., to downplay the influence of local maxima) is to increase the chance of moving from one arbitrary state to another.

As mentioned, however, given any choice of operators and energy metric, our hypothesis is that people improve when they are permitted to physically move letters because movement enables them to instantly “reset” their position in the landscape. Thus, when they find themselves trapped in a particular region, they can use tile rearrangement as an interactive strategy to assist internal search. In a sense, such physical reorganization provides an element of randomness that supports intelligent behavior (cf. Mitchell & Hofstadter, 1995).

We now turn to our experimental data on the Scrabble task.

Scrabble Experiment

The goal of the experiment was to examine performance in a word formation task using Scrabble letters. Specifically, we hypothesized that people would generate more words with a set of Scrabble tiles when allowed to physically ma-

nipulate the tiles than when not allowed to physically manipulate the tiles. Because we could not test the same person on the same set of tiles in both conditions, we first attempted to establish two sets of letters from which people naturally generate about the same number of words.

Norming Task

Six sets of seven letters each were created by randomly selecting tiles from the Scrabble game. Two sequences for each of the sets were randomly generated (e.g., “RDLOSNA” and “ARLNDOS”) to test whether there was an effect of the order of the letters.

Sixteen undergraduates from the University of California, Santa Cruz participated in the norming task to fulfill a requirement in a psychology course. Each participant saw one of the two sequences for each of the six originally chosen sets. The sequences and the order in which the sequences were presented were balanced across participants. Thus, eight participants saw one of the two sequences of each set.

In this pencil-and-paper task, participants were given five minutes to write down as many words as they could by rearranging the letters from each sequence. The participants were informed that words did not have to use all the letters in the sequence but could vary in length between two and seven letters.

For each of the twelve letter sequences, mean number of words generated was calculated. We then compared total number of words generated for each set. A series of *t*-tests between each pair of orders for a given set showed no effect for the number of words generated, so the results for each set were collapsed across the two orderings. A one-way analysis of variance (ANOVA) among the six sets of letters showed a significant effect, $F(1, 5) = 26.2$, $p < 0.001$, indicating that more words were generated for some of the sets than for others. Inspection of the data revealed that about the same number of words were generated for three of the six sets (see Table 1).

Table 1: Mean number of words generated per letter set.

Letter Sequence	Number of Words
“NDRBEOE”	19.88
“ESIFLCE”	12.06
“EMTGPEA”	22.25
“RDLOSNA”	20.81
“IRCDEOE”	16.19
“LNAOIET”	26.07

Another one-way ANOVA indicated that the number of words for “EMTGPEA”, “RDLOSNA”, and “NDRBEOE” did not differ, $F(1,2) = 1.01$, NS. Thus, we chose the first two of these three sequences as stimuli for our experiment because they shared the fewest letters.

Scrabble Method

Twenty undergraduates from the University of California, Santa Cruz participated in the experiment to meet a requirement in a psychology course. Each participant was a

native speaker of English or demonstrated a high proficiency in the language, as determined by responses on a questionnaire and vocabulary test given prior to the experiment.

The experiment was a 2 x 2 mixed design, with physical manipulation (Hands vs. No Hands) as the within-subjects factor, and letter sequence ("RDLOSNA" vs. "EMTGPEA") as the between-subjects factor. The letter sequence and the order in which the Hands or No Hands condition was performed was balanced across participants.

Participants were informed that they would have five minutes to generate as many English words as possible that were at least two letters long. Words were legal only if they were made from the tiles given. Participants were instructed not to use proper names (e.g., "Ron") or acronyms (e.g., "IBM"). They were also instructed to spell out the words as they found them (e.g., "TEAM, T-E-A-M") so that homophones (such as "BE" and "BEE") could be easily distinguished.

The task began with a practice trial. Half the participants were given instructions for the Hands condition, and the other half were given instructions for the No Hands condition. Participants in the Hands condition were told that they could use their hands to physically rearrange the tiles, but that it was not necessary to move the tiles to find and call out words. Participants in the No Hands condition were told that they could not use their hands to physically rearrange the tiles.

The set of Scrabble tiles was laid out on the table in front of the participant and the practice trial began. Practice proceeded for five minutes and then the participant performed a distractor task for five minutes. At this point, the participant performed the test trial in the same condition as the practice trial (i.e., Hands practice followed by Hands test). After five minutes of the word generation task, the participant moved to another distractor task, and then onto the other Hands or No Hands condition in the same way as before: practice followed by distractor followed by test. Throughout the practice and test trials, the experimenter transcribed the words as they were called out and the session was taped.

Results

The number of legitimate, unique words generated by each participant in each condition was calculated. The mean for the Hands condition was 20.70 (SD = 5.00) and for the No Hands condition, 19.30 (SD = 5.58). A two-way repeated measures ANOVA showed a main effect for the within-subjects factor (Hands vs. No Hands), $F(1, 18) = 5.165$, $p < 0.05$, indicating a difference in performance for Hands vs. No Hands. The effect of letter sequence was not significant, $F(1,18) < 1$, indicating there was no difference between "RDLOSNA" and "EMTGPEA". However, an interaction was obtained between manipulation condition and letter sequence, $F(1, 18) = 91.739$, $p < 0.0001$, indicating the use of hands had different effects on the two different sequences (see Table 2).

Table 2: Mean number of words generated.

	Hands	No Hands
"EMTGPEA" (n = 10)	23.30	16.00
"RDLOSNA" (n = 10)	18.10	22.60

To try to make sense of the manipulation by sequence interaction, post hoc comparisons were conducted. For "EMTGPEA", a significant difference was found between the Hands and the No Hands conditions, $t(18) = 4.97$, $p < 0.0001$; but for "RDLOSNA", no such difference was found, $t(18) = 1.87$, $p > 0.05$. Thus, physically moving the tiles improved performance for one letter sequence, but it had marginal and opposite effect on the other letter sequence.

Additional tests were conducted to investigate whether order of presentation (Hands, No Hands vs. No Hands, Hands) had an effect on the number words produced, or whether the average length of words produced varied across conditions. No effect for order was obtained, $t(18) < 1$, NS, indicating that the number of words generated did not depend on which condition (Hands or No Hands) was seen first. Similarly, there was no effect of manipulation condition or letter set on the average length of the words produced per participant, $F(1,18) < 1$, NS, in all cases. Overall, the mean word length was 3.30 (SD = 0.22).

Discussion

Overall, more words were generated when participants were allowed to manipulate the tiles than when they were not allowed to manipulate the tiles. This bears out our initial hypothesis. The interaction between manipulation and letter sequence, however, was unanticipated – the norming data were supposed to assure us that an equivalent number of words would be generated for both letter sequences.

One possible explanation for this interaction concerns the relative difficulty of producing words from the different letter sets. The more difficult it is to generate words from a set of letters, the more we suppose physical rearrangement would help. In terms of the state-space search model outlined previously, use of hands might be more effective if the words are more spread out in the space. In this case, physically rearranging the tiles has the effect of easily resetting the system's position in the energy landscape, enabling wider coverage during search.

One simple measure of word-generation difficulty might be word length. If the state-space search were based primarily on orthographic features and operations, we would expect longer words to be more difficult to generate, as long words must require more operations to compose than short words. Of course, we found no effect of the physical manipulation conditions on the length of words generated, suggesting that word length is not related to difficulty.

A semantic measure of word-generation difficulty relates to the productiveness of the letter strings and the frequency of the words that can be formed. First, 92 words can be generated from the letters "RDLOSNA", whereas only 53 can be generated from "EMTGPEA". Second, 47 of the 92 words in "RDLOSNA" do not appear in the Kucera and

Francis (1967) corpus of written English and the mean frequency of the remaining 45 words is 2735; nineteen of the 58 words in "EMTGPEA" do not appear in Kucera and Francis and the mean frequency of the remaining 39 is 336. In the Brown (1984) corpus of spoken English, the mean frequency for "RDLOSNA" is 1395, and for "EMTGPEA", 221. In both written and spoken English, the words contained in "RDLOSNA" are far more frequent in English than those contained in "EMTGPEA". Thus, it is plausible to suppose that physically arranging the letters would be helpful when trying to produce words from the less productive and less frequent set, as it must be more difficult to produce words in this case.

For the participants in the experiment who chose to use their hands very little or not at all, there was little difference between the Hands and No Hands conditions.² Interestingly, when asked why they did not move the letters more, these participants almost universally responded that they thought they could move the letters faster in their heads than they could on the table. As we have argued, this common intuition seems to be false in this case and in many others (Clark, 1997; Hutchins, 1995a, 1995b; Kirsh, 1995; Kirsh & Maglio, 1994).

Conclusion

We tested the hypothesis that physical actions can make problem solving easier. In our study, people were given sets of Scrabble letters and asked to generate words in two conditions: with their hands and without their hands. The results indicated that more words were generated when people used their hands than when they did not, although the story is somewhat more complicated. We argued that in this case, the physical actions of moving the letters allow people to effectively use the external environment as part of an interactive process of searching for words. In future experiments, we hope to control better for the productiveness of the strings and for the frequency of the words that can be produced from them.

Acknowledgments

Thanks to Ray Gibbs for supporting this research in his lab, and for many thoughtful comments and suggestions on this work. Thanks to Jenny Lederer for helping to run subjects. Thanks to Chris Campbell and Rom Brafman for useful discussions on experimental design and data analysis. Thanks to Chris Dryer, Denis Lalanne, and an anonymous reviewer for insightful comments on a draft of the paper.

References

- Brown, G. D. A. (1984). A frequency count of 190,000 words in the London-Lund corpus of English conversation. *Behavioural Research Methods Instrumentation and Computers*, 16, 502-532.
- Clark, A. (1997). *Being there: Putting body, brain, and world together again*. Cambridge, MA: MIT Press.
- Hofstadter, D. R. (1983). The architecture of JUMBO. In *Proceedings of the International Machine Learning Workshop* (pp. 161-170).
- Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science*, 19, 265-288.
- Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, 73, 31-68.
- Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior*, 4, 415-452.
- Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18, 513-549.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Maglio, P. P., & Kirsh, D. (1996). Epistemic action increases with skill. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 391-396). Mahwah, NJ: LEA.
- Mitchell, M., & Hofstadter, D. R. (1995). The Copycat project. In D. R. Hofstadter (Ed.) *Fluid concepts and creative analogies: Computer methods of the fundamental mechanisms of thought*. New York: Basic Books.

² Roughly one third of the participants in the Hands condition chose not to use their hands or used their hands only briefly. The small sample here makes statistical analysis difficult. We note this only as a passing observation.