Does ACT-R Model Me?

Emily Greve (emilygreve@cmail.carleton.ca)
Institute of Cognitive Science, 1125 Colonel By Dr.
Ottawa, ON. K1S 5B6 CA.

Elisabeth Reid (elisabethwood@cmail.carleton.ca)
Institute of Cognitive Science, 1125 Colonel By Dr.
Ottawa, ON. K1S 5B6 CA.

Robert West (robert.west@carleton.ca)
Institute of Cognitive Science, 1125 Colonel By Dr.
Ottawa, ON. K1S 5B6 CA.

Abstract

We were interested in testing Newell's Micro Strategies hypothesis as well as assumptions made by both ACT-R and SGOMS theory using a mobile game and a predictive SGOMS-ACT-R model. The Model is designed to predict expert game play. We found in most conditions the model did predict the results, however in one condition the player employed an alternative Micro Strategy.

Keywords: SGOMS; GOMS; ACT-R; Micro-cognition

Introduction

Often in studies on cognitive tasks many participants are used, and their results are averaged together to deal with variation in the results. This variation is generally interpreted as noise. However, Newell (1973) said that by averaging over many participants, we may be averaging over different strategies. If you control noise by training participants to the expert level, so they are on the same place on the learning curve, and rigorously control the task, different patterns in the data can be attributed to different Micro Strategies. Micro Strategies take place at the millisecond scale and can vary during simple cognitive tasks (Gray and Boehm-Davis, 2000).

In *You Can't Play 20 Questions with Nature and Win*, Newell (1973) notes the importance of understanding the different micro strategies for using our cognitive, perceptual, and motor systems to perform tasks in Cognitive Psychology experiments (note, Newell refers to these strategies as Methods, but we will use the term, Micro Strategies, and reserve the term, Methods, for use in our SGOMS model). For example, if a person needs to respond to a stimulus by typing a two-letter code, they could use one finger, one finger on each hand, or two fingers on the same hand if the letters were close together. If they know what code to type, they only need to see that a stimulus appears; they do not have to wait to fully recognize it before responding. However, some people might wait to fully register the identity of the stimulus before responding. Newell's point was that we should not average across different strategies as it produces meaningless numbers that do not accurately reflect the operation of the underlying cognitive system.

A previous study by West, Ward, Dudzik, Nagy, & Karimi (2018) used an ACT-R Agent (Anderson & Lebiere, 1998) built according to SGOMS (West & Nagy, 2007). The Agent they built was designed as a predictive model on expert gameplay. Only two participants were used but they were extensively trained to be high performing experts. The results of the study showed that two participants matched each other and the ACT-R Agent within milliseconds of accuracy under specific conditions, but not under others.

To further the research started by West et al. (2018), we decided to follow a similar experimental design but developed a version of the game without the conditions where the two subjects varied.

The Game

The new game, called Four Button was built using MIT app inventor 2 (https://appinventor.mit.edu/), and is run as an app on mobile devices. Four Button Expert levels follow the SGOMS structure and explicitly uses a hierarchy structure composed of operators (individual button presses), methods (fixed series of button presses), Unit Tasks, and Planning Units. Four Button Expert can be best explained by comparing gameplay to that of a First-Person Shooter videogame, such as. Dead Space, a videogame in which a player must fight through different levels of a game. Throughout the levels players encounter aliens which they must fight using different combinations of moves and weapons.

The Methods Level

The Methods level is equivalent to knowing which buttons correspond to which actions of the character. Buttons on a game controller such as X, O, and R2 correspond to actions such as Jump, duck, and shoot. In the Four Button Expert the Methods take the form of a two-letter prompt and a corresponding four-digit response. Players must enter the four numbers when prompted by the appearance of the two letters at the top of the screen (see Table 1). Expert players would have the four-number sequence proceduralized and be able to enter it immediately when they know which Method is required.

Table 1: The Methods

Methods	
AK	1234
SU	4123
ZB	2143
RP	4321
FJ	3214
HW	2341
YP	3412
WM	1432

The Unit Task Level

The Unit Task level is equivalent to knowing the different action sequences you must use to fight an enemy. For example if there is one alien type that you shoot until he executes/performs an attack on you upon which your you duck before resuming shooting, your button sequence would be R2, R2, O, R2. Compare this to another alien type where you must shoot, jump to avoid his attack type, shoot and then duck, in this case your button sequence would be R2, X, R2, O. Players must use a different sequence of the same actions in the different conditions. In our game the two-letter prompts are organized in specific and consistent sequences which then correspond to unit tasks (see Table 2). Expert players recognize that specific Methods signify the beginning of a Unit Task and know which related sequence of numbers are needed to complete the Unit Task. In two of the Unit Tasks (RP and HW) there are splits that occur, where one of two or one of three Methods could be displayed. The splits always occur at the same place in each Unit Task. This is the equivalent of some of the aliens having two or three different possible attack types that they employ at random at a certain point in the sequence.

Table 2. The Unit Tasks

Unit Tasks
Unit Task 1 (AK UT)
AK-WM-SU-ZB-FJ
Unit Task 2 (RP UT)
RP-SU< ZB-WM / YP-FJ
Unit Task 3 (HW UT)
HW-YP< FJ / ZB / SU

The Planning Unit Level

The Planning Unit level can be compared to a full level of our hypothetical video game. During a level, different aliens would be activated at different points throughout the level. An expert at this game would know that on level 1, Alien type 1 appears, followed by Alien type 2 and followed by Alien Type 3. Whereas on Level 2 the Alien order is Type 3, 1, then 2. Expert Players would be able to know exactly which order the Unit Tasks (Aliens) appear and which sequence of Methods (actions) take place within those. Our game follows this structure as well, where each Planning Unit holds the same Unit Tasks in different orders (see Table 3). Expert players are able to recognize which Planning Unit they are in by looking at the first Method code of the planning unit.

Table 3. The Planning Units

The Planning Units
Planning Unit 1
(AK UT)-(HW UT)-(RP UT)
Planning Unit 2
(RP UT)-(HW UT)-(AK UT)
Planning Unit 3
(HW UT)-(RP UT)-(AK UT)

Methodology

For our study we had 1 participant in order to thoroughly understand one individual's micro strategies before collecting more. Also, we could compare the results to similar conditions in the previous version of the game (West et al., 2018). This Methodology of making detailed comparisons between a few participants has also been successfully used by Gray and Boehm-Davis (2000), and Shiffrin and Cousineau (2004).

The game app was downloaded onto the participant's phone. The participant learned the game, starting at the Methods level. They moved up to the Unit Task level after they could confidently play each Method and their timing was consistent across all 8 Methods. Once they could play all three Unit Tasks they were moved up to the final level, the full game or Planning Unit level. Reaction time was based on how fast it takes for the participant to enter the corresponding four-digit code from when the two-letter code first appears on the screen.

Predictions

Based on the ACT-R model and previous results from West et al. (2018) we were able to make some predictions. For the conditions with multiple possible responses, we predicted that Hick's law, which states that reaction times increase as the number of stimulus–response alternatives increase (Hick 1952), would not apply. That is, we predicted no difference between our 3 choice and 2 choice conditions (see table 2). This is because the model assumes expert players will not rely on declarative memory (see Schneider and Anderson, 2011).

The model has one free parameter, which is the perceptual motor time to respond. However, this parameter is different depending on whether the next method could be memorized (known), or whether it was necessary to see the code before choosing the method (unknown). Because the known and unknown conditions have different estimated parameter values, we treated them separately. For the unknown condition the model predicts that a method signaling the beginning of a planning unit will take more time due to the SGOMS overhead required to keep track of the planning unit, compared to a method not associated with the beginning of a planning unit,. For the known condition, the model predicts that a method signaling the beginning of a unit task will take more time due to the SGOMS overhead required to keep track of the unit task, as opposed to a method not associated with the beginning of a unit task.

Finally, we predicted that our results should be the same as West et al. (2018) when scaled so that the known and unknown parameter values are the same across all subjects and the model.

Results

Methods where the player's data had errors were removed because we are interested in the player's time during conditions when they are playing correctly. To clean the data, it was sorted from smallest to largest time (in milliseconds). Outliers were cut off by detecting a knee in the data. (Satopaa, Albrecht, Irwin, & Raghavan, 2011). The mean average time was calculated from the remaining data in each condition.

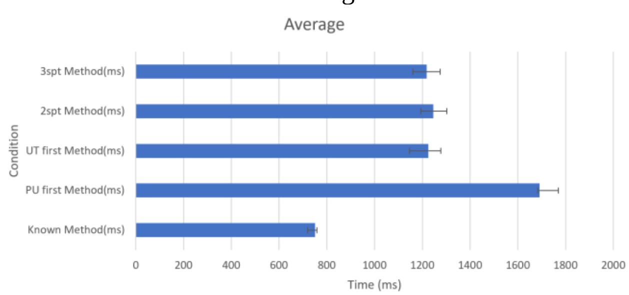

Figure 1: The Player Results (ms) under the different conditions with confidence intervals (0.05)

Figure 1 shows the player results with 0.05 confidence intervals. As predicted the three split (3spt method) and two split (2spt method) conditions were the same. Additionally, the predictions of extra processing time for unknown methods at the start of a planning unit (PU first method) and known methods at the start of a unit task (UT first method) was supported. This is illustrated by comparing PU first method to both 3spt method and 2spt method and comparing UT first method to known method.

Compared to the West et al. (2018) results, these results were scaled by assuming that differences in speed were due to perceptual/motor differences only. The model treats perceptual/motor as an additive factor, so we used the split conditions (which were combined into one condition) and the known methods condition to estimate the difference between participants in perceptual/motor speed for known perceptual motor actions and unknown perceptual motor actions. We then equalized perceptual/motor speed by subtracting an amount so that all participants were the same as the fastest participant in these two conditions, whose perceptual motor times were also used in the model. To test the model this same amount was also subtracted from the planning unit start condition (Pu first method) and the unit task start condition (UT first method), with the prediction that model and participants also be the same across these conditions.

The modeling results, displayed in Figure 2, show that our participant matched the model and the West et al. (2018) participants for the planning unit start condition, but took significantly more time for the unit task start condition. However, the time for the unit task start condition closely matched the unknown method time, suggesting that our participant cued off the displayed code rather than using their memory for this condition.

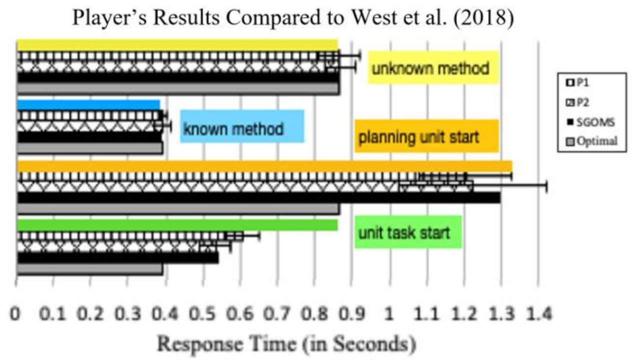

Figure 2. The data from our Participant compared to the result of the previous study. The hatched bars are the other two players from West et al. (2018), the black bar is the SGOMS/ACT-R model predictions and the gray bar is the optimal ACT-R prediction, where the model does not keep track of where it is in the task

Conclusion

Overall, we showed support for the idea that data and models can be used to study micro strategies in individuals. In particular, as in Gray and Boehm-Davis (2000) and Shiffrin and Cousineau (2004), we provide evidence that people can adopt different micro strategies even for simple tasks.

We can also make a prediction: If our analysis is correct, it should be possible to alter our participant's strategy by training them to rely on memory rather than vision for the unit task start condition. Such training should produce the predicted result. We will attempt this training and report the results at a future conference.

References

Anderson, J. R., & Lebiere, C. (1998). The atomic components of thought. Mahwah, NJ: Erlbaum.

Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of experimental psychology: Applied*, *6*(4), 322.

Hick WE. (1952) On the rate of gain of information. Quarterly Journal of Experimental Psychology. 4:11–26.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium.

Satopaa, V., Albrecht, J., Irwin, D., & Raghavan, B. (2011, June). Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops* (pp. 166-171). IEEE.

Schneider, D. W., & Anderson, J. R. (2011). A memory-based model of Hick's law. *Cognitive psychology*, *62*(3), 193–222.

Shiffrin, R., & Cousineau, D. (2004). Termination of a visual search with large display size effects. *Spatial vision*, *17*(4), 327-352.

West, R. L., & Nagy, G. (2007). Using GOMS for modeling routine tasks within complex sociotechnical systems: Connecting macrocognitive models to microcognition. *Journal of Cognitive Engineering and Decision Making*, *1*(2), 186-211.

West, R., Ward, L., Dudzik, K., Nagy, N., & Karimi, F. (2018, July). Micro and Macro Predictions: Using SGOMS to Predict Phone App Game Playing and Emergency Operations Centre Responses. In *International Conference on Engineering Psychology and Cognitive Ergonomics* (pp. 501-519). Springer, Cham.