Chapter 0

Dealing with TRACS: A Game of Chance and Choice

K. Burns

The discovery of chances and the management of choices require probabilistic inferences and dynamic investments. Here I present a card game called TRACS that was developed to study how, and how well, people perform these tasks. I compare the findings from human experiments to agent simulations in order to benchmark cognitive heuristics against normative strategies. I also discuss the practical implications for designing decision support systems.

1 Introduction

Human cognition is the central component in discovering *chances* and managing *choices* (Ohsawa 2003). Although every domain is different, the cognitive challenges of many domains can be boiled down to two tasks. One task is *inference*, which leads to some *confidence* in a situation; and the other task is *investment*, which leads to a *consequence* from some course of action.

Real world problems of chance discovery and choice management are complex because inferences must be made with uncertain (probabilistic) information and because investments must be made in temporal (dynamic) situations.

The question is: How, and how well, do people make these inferences and investments? Research on this question is best performed with *rigorous* control in the lab, but lab tasks must have *relevant* context if findings are to be of much use. Answers to the question can then be used to design systems that support human beings in practical problems of chance and choice.

Thus, with an eye towards both *rigor and relevance* in research on *chance and choice*, I developed a new family of card games for scientific investigations. The games, called TRACS (Tool for Research on Adaptive Cognitive Strategies, see Burns 2001), present players with tasks of probabilistic inference and dynamic investment that are similar (hence relevant) to problems in the world yet simplified (hence rigorous) for studies in the lab. In this chapter I describe the simplest game, played solitaire, and discuss the results of human experiments and agent simulations with this game.

1.1 The Deck

TRACS is a family of card games played with a special deck of two-sided cards (Burns 2001, 2004a). The backs of the cards show black shapes (triangle, circle, square) called *tracks*, and the fronts of the cards show colored sets (red, blue) called *treads*. The tracks provide probabilistic clues to the treads based on a non-uniform distribution of cards in the deck. The initial distribution (Table 1) defines the *baseline odds*, e.g., the red:blue odds are 6:2 for triangles, 4:4 for circles and 2:6 for squares. But this probabilistic structure changes in a dynamic fashion as tracks are turned to reveal their treads and then removed from play (like in Blackjack). The challenge is to count the cards (as they are turned) and update odds in order to make the best choice on each turn of the game. Here I consider only the simplest game, played solitaire, called Straight TRACS.

# of cards	6	4	2	2	4	6
track (back)	▲	●	■	▲	●	■
tread (front)	**red**	**red**	**red**	**blue**	**blue**	**blue**

Table 1. Distribution of track/tread cards in the deck.

1.2 The Game

Each turn in Straight TRACS is a choice between two tracks (shapes), where the goal is to turn the track that is most likely to match the tread (color) of a third card. The game goes like this: The deck is face down and three cards are dealt to a field. Two cards are dealt face down showing their tracks, and the third card is dealt face up in the middle showing its tread. The player turns over one of the two tracks to reveal its tread, trying to match the color of the middle card. The turn is scored as either a *save* (match) or *strike* (mismatch) and the two treads are removed from the field. The remaining track is turned to reveal its tread and this tread becomes the color to be matched on the next turn. Two new tracks are dealt from the deck, a track is turned and scored, etc. – until the deck is spent. The object is to make matches (avoid strikes).

1.3 The Test

The task of Straight TRACS, as described above, involves counting cards and updating odds in order to choose the best track on each turn. My test with this game was to see how well people could judge odds and choose cards, and how much they thought a support system would help. My probe was a confidence meter that measured a player's subjective belief in the red:blue odds for each track on the field (before the tracks are turned). The confidence meter resembled a colored ruler with buttons (equally spaced at eights) and a player clicked a button (on or between 100% red and 100% blue) to indicate his subjective belief.

2 Investigation

Figure 1 shows the normative problem and the cognitive data for a typical game and player, respectively. Tick marks on the y-axis correspond to buttons on the confidence meter. In the upper panel of Figure 1, the normative problem is that the red:blue odds change as cards are turned to reveal their treads. Notice the % red for triangles is the same or less than the % red for circles through most of the game. In the lower panel of Figure 1, the cognitive data for a typical player is biased towards the baseline (initial) odds. That is, the % red for triangles is always more than the % red for circles, which reflects the baseline relation between triangles and circles. [For squares, the % red is the same in upper and lower panels, as discussed in Section 2.2.].

2.1 Anchoring and Adjustment

The *baseline bias* (Burns 2002) seen in Figure 1 (lower panel) is consistent with a heuristic strategy of *anchoring and adjustment* (Tversky and Kahneman 1982) that has been observed in other studies of human judgment. That is, people are anchored to their initial assessment of the odds and make only minor adjustments thereto as the game proceeds. This is illustrated in Figure 2, which plots the experimental results obtained from 43 different people who each played 10 games of Straight TRACS (430 total games). The plot shows the average error in setting the confidence meter during the games. For comparison, Figure 2 also plots the error for a simulated agent that always sets the odds equal to the baseline odds. The comparison shows that human performance is similar to the baseline agent, although human performance is better than the baseline agent (less error) near the ends of games. Note that a perfect player who counts the cards and updates odds would be plotted as a horizontal line with zero error on each turn.

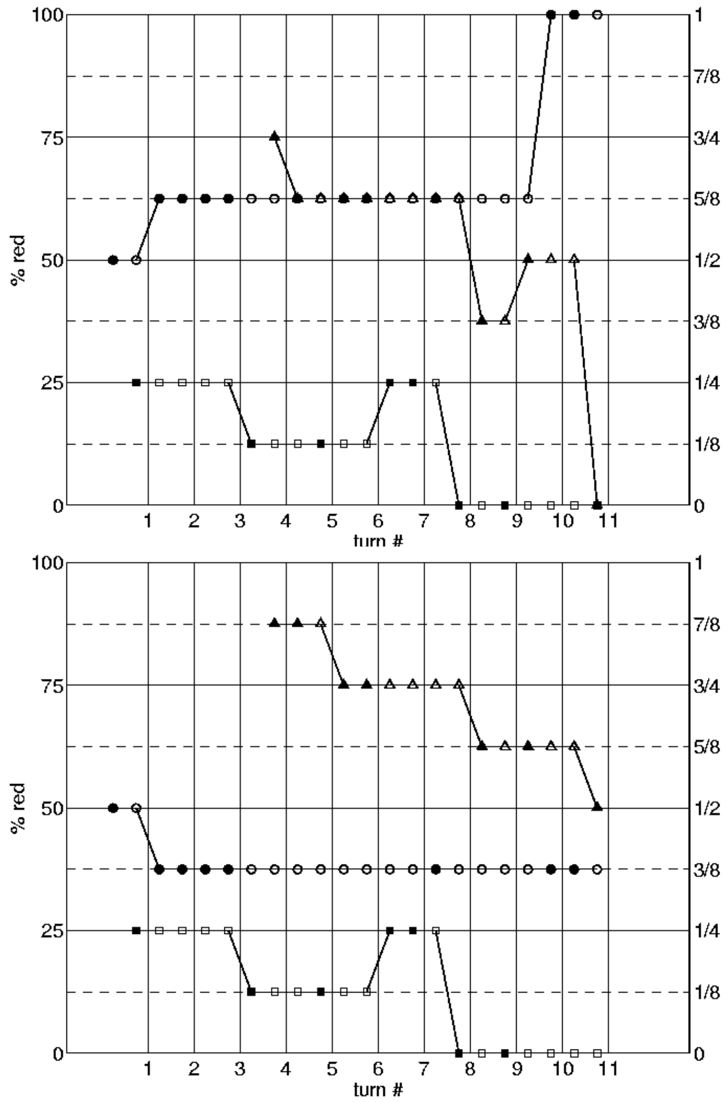

Figure 1. The normative problem (upper panel) and the cognitive data (lower panel) for a typical game and player. Compared to the objective odds (upper panel), the subjective odds (lower panel) are biased towards the baseline (initial) odds for triangles and circles (but not squares).

While human error follows the baseline error (Figure 2), it is important to note that players were not simply clicking the baseline odds (although they made many fewer adjustment than they should have). Figure 3 shows this by plotting the number of times that a player made an adjustment (from one turn to the next) of various magnitudes, where magnitude is measured by the span of buttons (separated by eights) on the confidence meter.

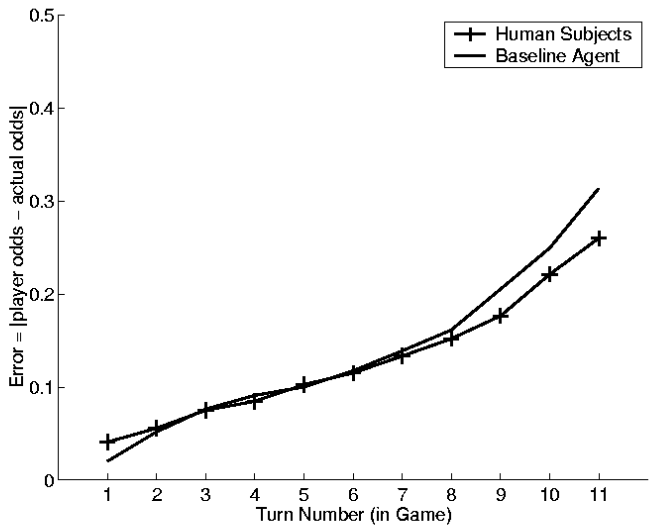

Figure 2. Average error versus turn, for human subjects and baseline agent.

2.2 Counting the Cards

The question is: What are the mental limits that prevent more accurate adjustments? This question involves two tasks, namely: *counting the cards* in each of six sets (see Table 1) and *normalizing the numbers* in order to convert the card count to a % chance. For example, given baseline odds of 6:2 (red:blue) = 75% red for triangles, and having seen 2 red triangles and 1 blue triangle, the updated (actual) odds for triangles are 4:1 = 80% red.

Although human errors arise in counting cards and normalizing numbers, I focused on the former task. My approach was to examine all instances where the normative odds reached 100% red or 100% blue (0% red), since then all N blue or red cards of a given type should have been counted.

There are six cards types, where N=2, 4 or 6 for each type (see Table 1). As a measure of performance, I took the total number of *hits* and *misses* (in all 430 games) that players made for each set (N=2, N=4 and N=6). A hit was scored whenever a player correctly selected the button for 100% red or 100% blue on the confidence meter. A miss was scored if a player selected another button when in fact all N of a given card type had been revealed.

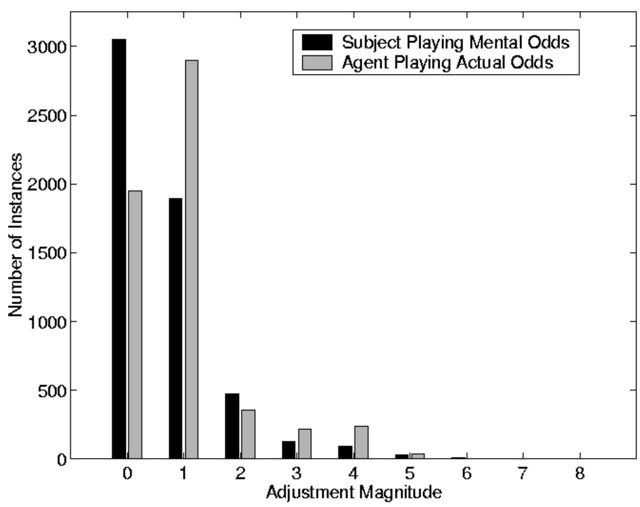

Figure 3. Number of adjustments made for each magnitude of adjustment.

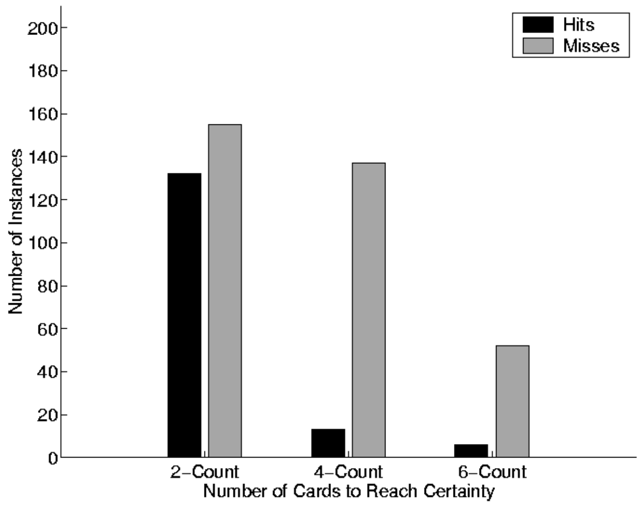

Figure 4. Cognitive competence in concurrent card counting.

Figure 4 shows that people are limited (<50% hit rate) in concurrent counting of all three sets (N=2, N=4, N=6), but there is a large drop in hit rate between N=2 and N=4. This suggests that a working limit on cognitive competence in concurrent counting is about 3±1 cards/set. Referring to Figure 1 (lower panel), note that the red squares (N=2) were counted correctly but the blue circles (N=4) and red triangles (N=6) were not counted correctly.

It is interesting to note that that a similar limit (3±1) has been measured for *subitizin*g in human visual perception (Dehaene 1992, Cowan 2001), which refers to how many things can be "seen all at once" (i.e., enumerated in parallel). It is also interesting to note that the standard method for card counting in Blackjack is to keep track of only one thing, namely the net count of high cards (10, J, Q, K, A) minus low cards (2, 3, 4, 5, 6) that have been seen (Mezrich 2002). TRACS was designed so that this method does not work, i.e., one must count cards in each of six sets (see Table 1) to judge odds that are useful in the game.

2.3 Illusion of Importance

With self-knowledge of one's limitations, an effective decision maker must *invest* internal (mental) and external (system) resources to achieve desired outcomes in "bounded rationality" (Simon 1981). As such, it is useful to examine how well people judge the benefits of various *satisficing strategies* that they might adopt and various *support systems* that they might acquire.

With regard to satisficing strategies, I polled players (post-test) to see which cards they thought were the most important ones to count (if they could only count some cards) in order to avoid strikes (mismatches) in Straight TRACS. Over half of all players thought the rare cards (red squares and blue triangles, see Figure 1) were the most important ones to count. Less than one quarter of all players thought the circles (red and blue) were the most important. But contrary to these beliefs, formal analyses and agent simulations showed that the circles were the most important cards to count and that counting only the rare cards had virtually no benefit.

With regard to support systems, I polled players (post-test) to see how much they thought a perfect card-counting system would help to avoid strikes (mismatches) in Straight TRACS. The players were asked how many strikes they would avoid, on average per game, if they had such a system. About 90% of all players said that they thought the system would allow them to save two or more strikes per game. This is a big benefit because the average number of strikes per game is typically about four. But contrary to this belief, formal analyses and agent simulations showed that the proposed system would save less than one strike per game (actually less than half a strike, on average per game). The reason, which players evidently did not appreciate (even after 10 games), is that the player is often faced with a "no win" situation where he is guaranteed to get a strike no matter which track he chooses.

These polls show how subjective beliefs can be biased by cognitive heuristics like *illusion of control* (Sage 1981). My work also shows how analyses and simulations can provide objective measures of just how good or bad cognitive heuristics really are in a specific context. For example, while *anchoring and adjustment* leads to serious limitations in estimating *confidence* (i.e., odds, see Figure 1), the ultimate performance of this heuristic in minimizing *consequences* (i.e., strikes, see above) is not much worse than perfect counting – at least in the game of Straight TRACS.

3 Implications

The bad news is that cognitive heuristics like *anchoring and adjustment* can lead to significant biases in subjective beliefs. The good news is that cognitive heuristics like *anchoring and adjustment* can be remarkably effective by some measures (e.g., consequences) even though they are extremely limited by other measures (e.g., confidence). The other bad news is that cognitive heuristics like *illusion of control* can also lead to significant biases in subjective beliefs about the benefits of support systems.

As such, the engineering challenge lies in developing and evaluating systems that can overcome the limits of cognitive heuristics while at the same time leveraging the power of cognitive heuristics. This is a balancing act that hinges on a fundamental understanding of human thinking, and TRACS research takes a step towards developing this understanding through human experiments and agent simulations. In this regard, TRACS is *relevant* because it simulates the probabilistic information and dynamic conditions found in *practical problems*; and TRACS is *rigorous* because it in facilitates the measuring and modeling of humans and agents in a *controlled context*. This blend of rigor and relevance is novel compared to previous research and useful for practical applications.

Further research on other TRACS games has shed light on other cognitive limits and led to the development of a prototype system for decision support in Bayesian *inference* (Burns 2004b, 2004c). Future research on TRACS is aimed at the two tasks of *inference* and *investment* in poker playing (Burns 2004a), since these two tasks are also the crux of *chance* discovery and *choice* management in practical domains (McDonald 1950).

References

Burns, K. (2001), "TRACS: A Tool for Research on Adaptive Cognitive Strategies", www.tracsgame.com.

Burns, K. (2002), "On Straight TRACS: A Baseline Bias from Mental Models", *Proceedings of the 24th Conference of the Cognitive Science Society*, pp. 154-159.

Burns, K. (2004a), "Making TRACS: The Diagrammatic Design of a Double-Sided Deck". In Blackwell, A., Marriott, K. and Shimojima, A., Eds., *Diagrammatic Representation and Inference*, Springer-Verlag, Berlin, pp. 341-343.

Burns, K. (2004b), "Bayesian Boxes: A Colored Calculator for Picturing Posteriors". In Blackwell, A., Marriott, K. and Shimojima, A., Eds., *Diagrammatic Representation and Inference*, Springer-Verlag, Berlin, pp. 382-384.

Burns, K. (2004c), "Painting Pictures to Augment Advice", *Proceedings of the International Working Conference on Advanced Visual Interfaces, AVI-2004,* Gallipoli, Italy, May 25-28.

Cowan, N. (2001), "The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity", *Behavioral and Brain Sciences*, vol. 24, pp. 87-185.

Dehaene, S. (1992), "Varieties of Numerical Abilities", *Cognition*, vol. 44, pp. 1-42.

McDonald, J. (1950), *Strategy in Poker, Business and War*, Norton, New York.

Mezrich, B. (2002), *Bringing Down the House: The Inside Story of Six MIT Students Who Took Vegas for Millions*, Free Press, New York, pp. 252-257.

Ohsawa, Y. (2003), "Modeling the Process of Chance Discovery". In Ohsawa, Y. and McBurney, P., Eds., *Chance Discovery*, Springer-Verlag, Berlin, pp. 2-15.

Sage, A. (1981), "Behavioral and Organizational Considerations in the Design of Information Systems and Processes for Planning and Decision Support", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 11(9), pp. 640-678.

Simon, H. (1981), *The Sciences of the Artificial*, MIT Press, Cambridge, MA.

Tversky, A. and Kahneman, D. (1982), "Judgment Under Uncertainty: Heuristics and Biases". In Kahneman, D., Slovic, P. and Tversky, A., Eds., *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, UK, pp. 3-20.