Review

Behavioural Isomorphism, Cognitive Economy and Recursive Thought in Non-Transitive Game Strategy

Benjamin J. Dyson [1,2,3]

1   Department of Psychology, P217 Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada;
    bjdyson@ualberta.ca
2   Department of Psychology, Ryerson University, Toronto, ON M5B 2K3, Canada
3   Department of Psychology, University of Sussex, Sussex BN1 9RH, UK

check for updates

Abstract: Game spaces in which an organism must repeatedly compete with an opponent for mutually exclusive outcomes are critical methodologies for understanding decision-making under pressure. In the non-transitive game rock, paper, scissors (RPS), the only technique that guarantees the lack of exploitation is to perform randomly in accordance with mixed-strategy. However, such behavior is thought to be outside bounded rationality and so decision-making can become deterministic, predictable, and ultimately exploitable. This review identifies similarities across economics, neuroscience, nonlinear dynamics, human, and animal cognition literatures, and provides a taxonomy of RPS strategy. RPS strategies are discussed in terms of (a) whether the relevant computations require sensitivity to item frequency, the cyclic relationships between responses, or the outcome of the previous trial, and (b) whether the strategy is framed around the self or other. The negative implication of this taxonomy is that despite the differences in cognitive economy and recursive thought, many of the identified strategies are behaviorally isomorphic. This makes it difficult to infer strategy from behavior. The positive implication is that this isomorphism can be used as a novel design feature in furthering our understanding of the attribution, agency, and acquisition of strategy in RPS and other game spaces.

Keywords: decision-making; behavioral isomorphism; cognitive economy; recursive thought; rock, paper, scissors

1. Competitive Decision-Making

There are a number of situations where an organism must repeatedly compete with others for mutually exclusive outcomes [1–3]): There will be only one Prime Minister or President, only one winner at Scrabble, only one bird able to forage nectar from any given flower. Within these domains encompassing " . . . cheetahs and gazelles, goalkeepers and penalty-kickers, cops and robbers . . . " ([4] p. 169), the maximization of gains and minimization of losses is essential for survival, and often requires domination of one's adversary. Adopting a level of behavioral predictability simpler than that of your opponent increases the possibility that, as prey you are captured, as a team you are eliminated, and as a deliverer of justice criminals remain elusive.

Game spaces have been established to study the dynamics of such competitive processes in a controlled fashion, where response and strategy are tractable. A number of spaces are available [5], but one space that continues to inspire research in economics [6], neuroscience [7], nonlinear dynamics [8], human [9,10], and primate [11,12] cognition, in addition to reflecting biological realities of the animal world itself (Uta stansburiana [13]; Drosophila melanogaster [14]), is that of rock, paper, scissors (RPS).

RPS serves as a three-response version of a simpler matching-pennies game (e.g., [15]). At each round of RPS, two players reveal a response from three options: Rock, paper, scissors, where rock beats (blunts) scissors, scissors beats (cuts) paper, and, paper beats (covers) rock. The empirical attractiveness of the game is derived from the relationships between the items [16], in that there is no one selection that is guaranteed to perform better than another. Due to the unique Nash equilibrium of the game [17], the only strategy that can guarantee the lack of exploitation is the mixed strategy or minimax solution, whereby all three responses are randomly played 33% of the time and without regard for the previous trial [18–21]. Approximations of the mixed strategy have been reported in pigeons [4], monkeys [22], and professional sports players [23,24]. This is in contrast to the position that mixed strategy is "cognitively extremely demanding and eventually implausible" ([25], p. 73) and the numerous empirical observations that reliably demonstrate the difficulties that individuals have attempting to behave randomly (e.g., [26–28]). Heterogeneity in the ability to express randomness is partly caused by the nature of the task. Commonly used 'production' tasks where participants are explicitly asked to behave randomly provide empirical evidence for deviations from randomness have been reviewed with respect to their logical and methodological problems [29]. In contrast, contexts where random performance is implicitly optimal such as competitive zero-sums game are more liberal tests of the ability to express mixed strategy [29]. Despite this, game performance deviates from stochastic behavior. To understand the varieties of predictable non-random performance on offer, an RPS strategic taxonomy is set out in terms of *frequency-*, *cycle-*, or *outcome*-based processes initiated from the perspective of the *self* or the *other*, based on empirical data from this game.

2. Taxonomy of Strategy in RPS

2.1. Frequency-Based Strategy

The first deviation from the mixed strategy is the overplay of one item (*frequency-based*). Specifically, the selection of rock currently enjoys a slight over-popularity in empirical studies of the game [8,10,16,19,30,31]. An increase in item selection frequency solely as a result of personal salience gives rise to the first type of RPS strategy: *Self-frequency*. In a collaborative environment where two organisms must select the same response, the use of *self-frequency* can allow individuals to gravitate towards joint selection, thereby maximizing their rewards [32]. However, in competitive environments, *self-frequency* strategies can result in exploitation. As [33] observe, if a computer opponent plays one item more often than another (e.g., rock) then human participants will play the appropriate counter-item with increased frequency (e.g., paper). Therefore, an increase in item selection frequency as a result of identifying primary salience in one's opponent (secondary salience; [32]) may be referred to as *other-frequency*, and can enjoy temporary dominance over an opponent whose own frequency distribution is biased.

2.2. Cycle-Based Strategy

The second set of strategies available during RPS result from the *non-transitive dominance relations* [16] between the items (henceforth, *cycle-based*). As shown in Figure 1, item selection can change across consecutive encounters in one of two directions. Figure 1a shows the selection of an item at trial *n* + 1 that would have beaten the previous item at trial *n* (e.g., rock followed by paper), whereas Figure 1b shows the selection of an item in trial *n* + 1 that would have been beaten by the previous item at trial *n* (e.g., rock followed by scissors). In the first instance, such behavior has been described as an 'ascending' [19], 'right-shift' [34], or 'one-ahead' [35] strategy, whereas in the second instance, behavior has been described as a 'descending' [19] or 'left-shift' [34] strategy. *Upgrade* and *downgrade*, respectively, will be used for the reminder of the paper. Moreover, due to the cyclical nature of the relationships between items, it is possible to repeat the strategy of *upgrading* or *downgrading* across multiple consecutive trials, and there is evidence that individual do this, at least in the short term [10].
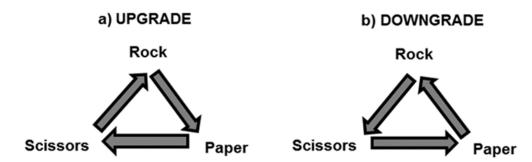
Figure 1. Recursive (**a**) upgrade and (**b**) downgrade item change strategies in rock, paper, scissors.

It is again important to consider whether the updating of response is primarily driven by one's own (*self-cycle*) or one's opponent (*other-cycle*) behavior. In the same manner that it is sub-optimal to maintain a single response during play (*self-repeat*; [4]), *self-cycle* behavior would also appear potentially myopic and easily dominated by an attentive opponent. On the other hand, other-item strategies are synonymous with Cournot dynamics [36], where Cournot's best response (CBR) represents the selection of an item that would have beaten the opponent's previous play (*other-upgrade*; the behavioral preference of schizophrenics, [19]). Cournot's second-best response (CSBR) is the selection of an item that would have drawn with the opponent's previous play (*other-repeat*; also referred to as 'one-back' by [35], and Cournot's worst response (CWR) is the selection of an item that would have been beaten by the opponent's previous play (*other-downgrade*).

2.3. Outcome-Based Strategy

What is critical to the current review is *other-cycle* strategies (see above) are behaviorally equivalent to the implementation of *self-outcome* strategies. The traditional mechanics of *outcome-based* strategy are rooted in model-free reinforcement learning principles where the outcome of an event influences the subsequent weighting of future responding [37]. Stemming from behaviorist laws like the *Law of Effect* [38], common expressions of such strategies may be summarized by the joint principles of *win-stay/lose-shift* (e.g., [12,16]). In the context of *self-outcome* strategy, a previous response will be more likely to be repeated as a result of winning, and changed as a result of losing. From an evolutionary point of view, failure to initiate behavioral change following a *loss* (or even the perceived threat of *loss*) is likely to be more damaging than the failure to repeat an action following a *win*. To wit: "Neither the mouse nor the gazelle can afford to *learn* to avoid" ([39], p. 33, emphasis in original). Some of the most recent work on RPS has been interested in the *direction* of shift following negative outcomes [16], thereby incorporating aspects of *cycle-based* strategy into *self-outcome* strategy [10,30].

3. Behavioral Isomorphism in RPS Strategy

Despite this seemingly large array of strategies available during RPS play, many are simply re-descriptions of the same mechanics, with different assumptions regarding the stance of the organism. Of primary note is the behavioral equivalence of an *other-cycle* (*other-upgrade* or CBR) strategy and the implementation of a traditional *self-outcome* (*win-stay/lose-shift*) strategy. To take the first line in Table 1 as a concrete example, a player winning as a result of paper should be more likely to repeat the paper response on the next trial (*win-stay*), which is equivalent to upgrading the opponent's previous play of rock also to paper (*other-upgrade*). All other eight possible outcomes follow the same pattern: The heuristic set out by the standard version of the *win-stay/lose-shift* self-outcome strategy yields the same trial *n* + 1 behavior as does simply utilizing an *other-upgrade* heuristic.

Furthermore, given the isomorphism between *other-cycle* and *self-outcome* strategies more generally, it is easy to go on to demonstrate that a simplistic *other-repeat* strategy (CSBR) can also be viewed as a non-standard revision of the self-outcome strategy where repeats are associated with draws, downgrades associated with wins, and upgrades associated with losses (Supplementary Table S1),

and an *other-downgrade* strategy (CWR) representing a similarly revised self-outcome strategy where repeats are associated with losses, downgrades associated with draws, and upgrades associated with wins (Supplementary Table S2). Such an expansion of this logic is not without biological precedent. For example, *win-shift* rather than *win-stay* strategies have utility for species foraging in environments with a fast depletion rate [40]. For example, nectarivorous birds are attuned to *win-shift* behavior as a result of the one-shot exploitation of flower nectar, but retain behavioral adaptation such that they can learn counter contingencies like *win-stay* if their environment changes [41]. Therefore, while seemingly counter to traditional reinforcement learning principles, the observation of *win-shift* strategies may be influenced by both the demands of the environment and the degree of species intelligence [40,42].

Table 1. Equivalence of deploying an other-upgrade strategy (also Cournot's best response) and a traditional self-outcome strategy.

Trial n		Strategy	Trial $n + 1$	Trial n	Strategy	Trial $n + 1$
Other	**Self**	*Other-cycle*	**Self**	**Outcome**	*Self-Outcome*	**Self**
Rock	Paper	UPGRADE	**Paper**	Win	WIN-REPEAT	**Paper**
Rock	Scissors	UPGRADE	**Paper**	Lose	LOSE-DOWNGRADE	**Paper**
Rock	Rock	UPGRADE	**Paper**	Draw	DRAW-UPGRADE	**Paper**
Paper	Scissors	UPGRADE	**Scissors**	Win	WIN-REPEAT	**Scissors**
Paper	Rock	UPGRADE	**Scissors**	Lose	LOSE-DOWNGRADE	**Scissors**
Paper	Paper	UPGRADE	**Scissors**	Draw	DRAW-UPGRADE	**Scissors**
Scissors	Rock	UPGRADE	**Rock**	Win	WIN-REPEAT	**Rock**
Scissors	Paper	UPGRADE	**Rock**	Lose	LOSE-DOWNGRADE	**Rock**
Scissors	Scissors	UPGRADE	**Rock**	Draw	DRAW-UPGRADE	**Rock**

4. Differences in Cognitive Economy

While there is behavioral isomorphism between *self-outcome* and *other-cycle* strategies, these approaches clearly differ in their perceived processing demands and the perspective from which each strategy is calculated (see also [43], and their discussion of cognitive feasibility in the comparison of contingent average and discounted average rule systems). Not only do organisms often enter competition with incomplete information regarding their opponent's rule structure [43], but they are also bounded by the limits of their cognition when trying to retain such information. As [44] note, working memory plays a critical role in the updating of strategy during collaborative and competitive environments, and to be successfully implemented, strategic demands cannot exceed working memory capacity limits (see also [7,45]). From an information processing point-of-view, there is cognitive economy for the organism in computing a response based on any *other-cycle* strategy as opposed to any *self-outcome* strategy. This is because the amount of information and the number of rules required to instantiate behavior on the basis of the *other-cycle* heuristic is less than the amount of information and the number of rules required to instantiate behavior on the basis of the *self-outcome* heuristic. To formalize the computations required by both approaches, the *other-upgrade* strategy detailed in Table 1 may be encapsulated by three conditional rules (Formulae (1)–(3); where O = *other*, S = *self*, r = rock, p = paper, s = scissors, n = current trial, $n + 1$ = subsequent trial):

$$\text{IF } O(n) = r \quad \text{THEN} \quad S(n + 1) = p \tag{1}$$

$$\text{IF } O(n) = p \quad \text{THEN} \quad S(n + 1) = s \tag{2}$$

$$\text{IF } O(n) = s \quad \text{THEN} \quad S(n + 1) = r. \tag{3}$$

In contrast, resolution of the standard *self-outcome* strategy (Table 1) requires a larger set of more complex conditionals, where the self-item from the previous trial and the outcome of that trial interact (Formulae (4)–(12); where, additionally, W = win, L = lose, D = draw):

$$\text{IF } S(n) = p \quad \text{AND} \quad (n) = W \quad \text{THEN} \quad S(n + 1) = p \tag{4}$$

$$\text{IF } S(n) = s \quad \text{AND} \quad (n) = L \quad \text{THEN} \quad S(n + 1) = p \tag{5}$$

$$\text{IF } S(n) = r \quad \text{AND} \quad (n) = D \quad \text{THEN} \quad S(n + 1) = p \tag{6}$$

$$\text{IF } S(n) = s \quad \text{AND} \quad (n) = W \quad \text{THEN} \quad S(n + 1) = s \tag{7}$$

$$\text{IF } S(n) = r \quad \text{AND} \quad (n) = L \quad \text{THEN} \quad S(n + 1) = s \tag{8}$$

$$\text{IF } S(n) = p \quad \text{AND} \quad (n) = D \quad \text{THEN} \quad S(n + 1) = s \tag{9}$$

$$\text{IF } S(n) = r \quad \text{AND} \quad (n) = W \quad \text{THEN} \quad S(n + 1) = r \tag{10}$$

$$\text{IF } S(n) = p \quad \text{AND} \quad (n) = L \quad \text{THEN} \quad S(n + 1) = r \tag{11}$$

$$\text{IF } S(n) = s \quad \text{AND} \quad (n) = D \quad \text{THEN} \quad S(n + 1) = r. \tag{12}$$

Both mechanics arrive at the same behavioral outcomes, but the *other-upgrade* strategy is outwardly and somewhat aggressively framed in terms of its exclusive focus on the opponent, while the *self-outcome* strategy is inwardly framed in terms of its more complex focus on the interaction between the previous item played by the organism and previous trial outcome. The increased working memory demands of the *self-outcome* strategy are similarly highlighted by [46] in the context of *Columba livia* learning: "to use a win-stay/lose-shift rule, the pigeon must remember not only the outcome of its last response but also the stimulus alternative to which it most recently responded." (p. 65). While *self-outcome* logic may be beyond the computational scope of certain organisms, the implementation of *other-item* logic is also problematic in terms of the stance the organism required to take. Despite its cognitive

economy, the difficulty in accepting the description associated with *other-item* strategies is that the mechanisms appear to ignore personal reinforcement and environmental history. As [47] further emphasize, successful performance in a wide range of behavior such as foraging, gambling, and investing all depend on the modelling of rewarded and unrewarded outcome distributions.

On the basis of a simplistic behavioral analysis, it is impossible to ascertain whether the organism is operating on the basis of the previous actions of the self or an opponent. To compound the issue, the distinction between *self-* and *other*-orientated strategies highlight a further descriptive fissure in terms of whether it is necessary to invoke recursive thought to account for RPS performance.

5. Differences in Recursive Thought

Since success in competitive environments rely on the outmaneuvering of one's opponent either by virtue of behavioral unpredictability [4] or by exploiting a level of opponent predictability simpler than one's own, a certain level of recursive thought appears necessary to 'second-guess' or think 'one-step-ahead' of one's competitor [3,48]. An organism's slavish sensitivity to its own reinforcement history (*self-outcome*) may prevent it from learning fast when there is complex environmental change [49] and there is evidence that despite its demands on working memory, the use of *win-stay/lose-shift* strategy in the context of human probability matching is actually linked to *reduced* individual estimates of working memory [50]. Furthermore, strategies such as those modelled on self-outcome contingencies remain predictable and hence exploitable if competitors are more sophisticated in their reasoning. For example, [51] document the use of *win-stay/lose-shift* behavior in football match squad configurations across consecutive games, but additionally note that such behavior was unsuccessful in influencing future game outcome.

Unfortunately, the logic of recursion in RPS again fails to distinguish between lower- and higher-order cognition. To demonstrate this, consider a case beyond a *self-outcome* strategy, where recursive thinking is introduced on the basis of the opponent. Here, not only does the player prepare an initial self-response based on the outcome of the previous trial, but they correct their initial response on the basis of what they assume their opponent would do to counter their initial response (a *second-order self-outcome* conditional; see also [27] p. 236). Note that this does not necessarily entail the appeal to theory-of-mind however, as reasoning about opponents' behavioral strategy is not synonymous with reasoning about what the opponent may be thinking [52,53]. Taking the top line of Table 2 as a concrete example, the player plans to repeat their paper response following a *win* against an opponent's rock. If the player thinks their opponent is sensitive to their *win-stay* strategy, their opponent should counter with scissors (beating their expected paper), meaning that the ultimate response of the participant should be rock (beating their opponent's expected Scissors). *Self-outcome* logic can be extended with the addition of {} to represent levels of recursion:

$$
\begin{aligned}
\text{IF } S(n) = p \quad &\text{AND} \quad (n) = W \quad \text{THEN} \quad S(n+1) = p \\
\{\text{IF } S(n+1) = p \quad &\text{THEN} \quad O(n+1) = s, \\
\text{IF } O(n+1) = s \quad &\text{THEN} \quad S(n+1) = r\}.
\end{aligned} \tag{13}
$$

The behavioral outcomes derived from applying a *second-order self-outcome* conditional are identical to Supplementary Table S1, which maps out a simpler, non-recursive *other-repeat* (also a variant of *self-outcome*) strategy. Consequently, the simple examination of behavior once again fails to distinguish between an organism engaged in second-*order* recursive thought focused on the self in terms of item and outcome, and an organism simply copying the response of its combatant seemingly without regard for the outcome of its own play. On the one hand, the demonstration that RPS behavior fails to distinguish between complex recursive thought involving multiple iterations of counter-play, relatively simple *outcome* conditionals and simpler still *cycle* strategies, is problematic when the goal is to interpret strategy from behavior. However, this behavioral isomorphism becomes a useful empirical tool with which to explore the mechanisms of strategy *attribution*, *agency,* and *acquisition*.

Table 2. Logic of a *second-order self-outcome* conditional.

Trial *n*			Strategy	Trial *n* + 1	Trial *n* + 1	Trial *n* + 1
Other	**Self**	**Outcome**	*Self-Outcome*	**Self**	**Other**	**Self (*Revised*)**
Rock	Paper	Win	WIN-REPEAT	Paper	Scissors	**Rock**
Rock	Scissors	Lose	LOSE-DOWNGRADE	Paper	Scissors	**Rock**
Rock	Rock	Draw	DRAW-UPGRADE	Paper	Scissors	**Rock**
Paper	Scissors	Win	WIN-REPEAT	Scissors	Rock	**Paper**
Paper	Rock	Lose	LOSE-DOWNGRADE	Scissors	Rock	**Paper**
Paper	Paper	Draw	DRAW-UPGRADE	Scissors	Rock	**Paper**
Scissors	Rock	Win	WIN-REPEAT	Rock	Paper	**Scissors**
Scissors	Paper	Lose	LOSE-DOWNGRADE	Rock	Paper	**Scissors**
Scissors	Scissors	Draw	DRAW-UPGRADE	Rock	Paper	**Scissors**

6. Future Work into Attribution, Agency, and Acquisition

Attributions regarding opponency are critical in determining the quality of decision-making. For example, [54] argue that evidence for primates making inferences about others may depend on the competitive nature of the environment, whereas [55] showed that opponents who demonstrated less sensitivity to previous outcomes may also encourage similar insensitivities in the competing organism. Therefore, irrational decisions defined as deviations from mixed strategy may be more likely in non-threatening environments, such as where primates can remain behaviorally vulnerable if their opponent fails to take advantage of strategic weakness [22]. One problem though in assessing an organism's sensitivity to potential competitive threat via the adjustment of opponent strategy is that changes in the distribution of items and outcomes may introduce additional confounds in working memory demands [44]. However, the empirical use of differentially described but behaviorally isomorphic RPS strategy help to resolve this issue. For example, identical opponent behavior could be described to the participant in two different ways across separate blocks: "the computer's strategy will be based on your previous selection" vs. "the computer's strategy will be based on its last selection and outcome" ([56], p. 1484). Here, the attempt is to test the impact of perceiving an opponent either as seemingly aggressively focused on the player (i.e., *other-upgrade*) or an opponent seemingly solipsistically focused on its own responses and outcomes (i.e., *self-outcome*; see the above section on *Differences in cognitive economy*). Thus, the impact of perceived increased (*other-upgrade*) or decreased (*self-outcome*) competitive threat on behavior could be examined without recourse to actual behavioral change between the opponents. Such experiments represent manipulations of language only, thanks to the behavioral isomorphism between certain RPS strategies. Data such as these could help to reveal the relative fitness of the organism in terms of responding and adapting to variable levels of competitive threat from both first-hand (experience) and second-hand (description) environmental information [57].

A second research strand that opens up is the study of how *agency* influences competitive decision-making. For example, [58] report an increased willingness for human participants to accept inequitable offers during an ultimatum game when their opponent was thought to be a computer relative to another human. Such a perceived lack of agency in the case of computerized opponency also appears linked to a reduction in skin conductance modulation [59], suggestive of decreased emotional reactivity to wins and losses when interacting with known automaton. Data from other game spaces suggest that the experience of negative emotion plays a key role in determining relatively poor-quality decision-making. In the case of RPS, the higher subjective value of losing relative to winning is a likely driving force behind the increased predictability of *lose-shift* behavior [10,30]. Similarly, the phenomenon of tilting in the poker community [60,61] describes a deterioration in the quality of decision making following loss including chasing behavior, wherein players attempt to recoup lost bets often by increasing stakes (c.f., Martingale strategy in roulette; [62]). Therefore, belief in competitive interaction with automata relative to conspecifics should reduce the emotional experience of loss and allow for rational rather than irrational subsequent behavior. Additionally, identifying those individuals for whom automated interactions yield equivalent or even heightened emotional response will be critical for understanding the antecedents of pathological gambling [63]. For example, fixed-odds betting terminals (FOBTs) are both automated and completely random in their behavior, thereby offering no recourse for strategy. Nevertheless, gamblers interacting with machines often ascribe agency to them, such as being more likely to play machines that have not recently paid out [42,64]. Compounded with their variable schedules of reinforcement [65], they represent a major temptation for repetitive gambling behavior without the need for interpersonal interaction. This understanding of expectations from both sentient and programmed agents is crucial as more and more virtual interactions are developed where the boundaries between human and automated respondents are increasingly blurred. The manipulation of perceived opponent agency against a backdrop of behavioral isomorphism then becomes a robust paradigm within which these questions can be addressed. For example, one might imagine an experiment where RPS is delivered across

remote computer terminals. Participants believe that their opponents alternate between a variety of human players and automated players but some of the behavioral patterns are exact between the two categories. This might provide additional insights regarding reactions to, and consequences of, success and failure against humans vs. computers [58,59], again with keeping all opponent behavioral parameters equivalent.

Finally, differences in internally (*self-*) or externally (*other-*)focused strategy may have varying degrees of salience for individuals impacted by Autistic Spectrum Disorder (ASD). As such, the presentation of opponent strategy may have significant consequences for the successful *acquisition* of environmental contingencies and the ability to improve performance in competitive domains. ASD is characterized by difficulties in the internal regulation of performance [66] and the association of reward with social stimuli [67]. If individuals on the autistic spectrum differ their ability to learn and subsequently dominate opponent strategy depending on the perspective of that strategy, then such an approach offers significant promise for the improvement of higher-order cognition in special populations. Specifically, if the same behavior can be described in both a social (*other-*) and non-social (*self-*)way, then individual with ASD might show faster and more reliable acquisition of learning when the description of their opponent's actions is more consistently framed within their own cognitive style (i.e., non-social or self-focused).

Despite its relative simplicity and ubiquity, RPS yields complex and potentially recursive patterns of data that can compromise an organism's dominant position during competition. Deviations from mixed strategy have been described with respect to sensitivity to item *frequency*, the *cyclic* relationships between responses, or the *outcome* of the previous trial, and whether the strategy is described in terms of the *self* or the *other*. One negative implication from this review is that despite differences in cognitive economy and the potential recruitment of recursive thought, many of these strategies have behavioral isomorphism and so caution is warranted when inferring strategy from behavior. However, the more positive interpretation of this conclusion is that such behavioral isomorphism introduces fruitful avenues of research into strategic attribution, agency, and acquisition that can be applied not only to RPS, but also to other competitive game spaces in which individuals compete for mutually exclusive outcomes.

In addition to the general caution raised by the current review in terms of attributing either relatively low-level or high-level cognition to organisms on the basis of their competitive behavior alone, the specific analysis of RPS also yields unique insights as a result of its novel structure. One key feature of competitive experience in the real world—and in contrast to similar encounters developed in the laboratory—is that the information we have about environmental contingencies and the outcome of our actions is often ambiguous [68]. As such, the presence of draw trials in a standard three-response, three-outcome version of RPS provides an example of an ambiguous outcome that may be compared against the consequences of more transparent *wins* and *losses* (e.g., [68–70]). Reactions to draws are of interest as they may be interpreted as either positive or negative [69,71,72], and this interpretation may impact on the overall sense of success felt during the game. As ([27], p. 225) state: "to tie is to fail to win, but on the other hand to tie is to avoid a loss." Interestingly, neural activity following wins, losses, and draws (feedback-related negativity, FRN; [73]) suggests that ambiguous outcomes generate a response statistically different from wins but statistically indistinguishable from losses [69]. Thus, games containing draw trials might weight the distribution of positive and negative outcomes more towards the anticipation of goal-failure rather than goal-success, potentially accounting for the preponderance of shift behavior relative to stay behavior, simply because negative outcomes are more frequent than positive outcomes. Behavioral and neural comparisons of RPS with a simpler two-response, two-outcome game such as matching pennies (MP; [15,31]) will help to explain how the presence of ambiguous outcomes contribute to behavior during competitive environments, and this is the focus of our current empirical work.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Decety, J.; Jackson, P.L.; Sommerville, J.A.; Chaminade, T.; Meltzoff, A.N. The neural basis of cooperation and competition. *NeuroImage* **2004**, *23*, 744–751. [CrossRef] [PubMed]
2. Goodie, A.S.; Doshi, P.; Young, D.L. Levels of theory-of-mind reasoning in competitive games. *J. Behav. Decis. Mak.* **2012**, *25*, 95–108. [CrossRef]
3. Yoshida, W.; Dolan, R.J.; Friston, K.L. Game theory of mind. *PLoS Comput. Biol.* **2008**, *4*, e10000254. [CrossRef] [PubMed]
4. Sanabria, F.; Thrailkill, E. Pigeons (Columba livia) approach Nash equilibrium in experimental matching pennies competition. *J. Exp. Anal. Behav.* **2009**, *91*, 169–183. [CrossRef]
5. Coleman, A.M. Cooperation, psychological game theory, and limitation of rationality in social interaction. *Behav. Brain Sci.* **2003**, *26*, 139–153. [CrossRef]
6. Xu, B.; Zhou, H.-J.; Wang, Z. Cycle frequency in standard Rock-Paper-Scissors games: Evidence from experimental economics. *Phys. A* **2013**, *392*, 4997–5005. [CrossRef]
7. Gallagher, H.L.; Jack, A.I.; Roepstorff, A.; Frith, C.D. Imaging the intentional stance in a competitive game. *NeuroImage* **2002**, *16*, 814–821. [CrossRef]
8. Toupo, D.F.P.; Strogatz, S.H. Nonlinear dynamics of the rock-paper-scissors game with mutations. *Phys. Rev.* **2015**, *91*, 052907. [CrossRef]
9. Cook, R.; Bird, G.; Lünser, G.; Huck, S.; Heyes, C. Automatic imitation in a strategic context: Players of rock-paper−scissors imitate opponents' gestures. *Proc. R. Soc. B Biol. Sci.* **2012**, *1729*, 780–786. [CrossRef]
10. Dyson, B.J.; Wilbiks, J.M.P.; Sandhu, R.; Papanicolaou, G.; Lintag, J. Negative outcomes evoke cyclic irrational decisions in Rock, Paper, Scissors. *Sci. Rep.* **2016**, *6*, 20479. [CrossRef]
11. Gao, J.; Su, Y.; Tomonaga, M.; Matsuzawa, T. Learning the rules of the rock-paper-scissors game: Chimpanzees versus children. *Primate*, in press.
12. Lee, D.; Conroy, M.L.; McGreevy, B.P.; Barraclough, D.J. Reinforcement learning and decision making in monkeys during a competitive game. *Cogn. Brain Res.* **2004**, *22*, 45–58. [CrossRef]
13. Sinervo, B.; Lively, C.M. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **1997**, *380*, 240–243. [CrossRef]
14. Zhang, R.; Clark, A.G.; Fiumera, A.C. Natural genetic variation in male reproductive genes contributes to non-transitivity of sperm competitive ability in *Drosophila melanogaster*. *Mol. Ecol.* **2013**, *22*, 1400–1415. [CrossRef]
15. Belot, M.; Crawford, V.P.; Heyes, C. Players of matching pennies automatically imitate opponents' gestures against strong incentives. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2763–2768. [CrossRef]
16. Wang, Z.; Xu, B.; Zhou, H.-J. Social cycling and conditional responses in the Rock-Paper-Scissors game. *Sci. Rep.* **2014**, *4*, 5830. [CrossRef]
17. Nash, J. Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. USA* **1950**, *36*, 48–49. [CrossRef]
18. Abe, H.; Lee, D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* **2011**, *70*, 731–741. [CrossRef]
19. Baek, K.; Kim, Y.-T.; Kim, M.; Choi, Y.; Lee, M.; Lee, K.; Hahn, S.; Jeong, J. Response randomization of one-and two-person Rock-Paper-Scissors games in individuals with schizophrenia. *Psychiatry Res.* **2013**, *207*, 158–163. [CrossRef]
20. Bi, Z.; Zhou, H.-J. Optimal cooperation-trap strategies for the iterated rock-paper-scissors game. *PLoS ONE* **2014**, *9*, e111278. [CrossRef]
21. Zhou, H.-J. The rock-paper-scissors game. *Contemp. Phys.* **2016**. [CrossRef]

22. Lee, D.; McGreevy, B.P.; Barraclough, D.J. Learning decision making in monkeys during a rock-paper-scissors game. *Cogn. Brain Res.* **2005**, *25*, 416–430. [CrossRef]

23. Palacios-Huerta, I. Professional play minimax. *Rev. Econ. Stud.* **2003**, *70*, 395–415. [CrossRef]

24. Walker, M.; Wooders, J. Minimax play at Wimbledon. *Am. Econ. Rev.* **2001**, *91*, 1521–1538. [CrossRef]

25. Griessinger, T.; Coricelli, G. The neuroeconomics of strategic interaction. *Curr. Opin. Behav. Sci.* **2015**, *3*, 73–79. [CrossRef]

26. Neuringer, A. Can people behave "randomly"? The role of feedback. *J. Exp. Psychol. Gen.* **1986**, *115*, 62–75. [CrossRef]

27. West, R.L.; Lebiere, C. Simple games as dynamic, coupled systems: Randomness and other emergent properties. *Cogn. Syst. Res.* **2001**, *1*, 221–239. [CrossRef]

28. West, R.L.; Lebiere, C.; Bothell, D.J. Cognitive architectures, game playing, and human evolution. In *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*; Sun, R., Ed.; Cambridge University Press: Cambridge, UK, 2006; pp. 103–123.

29. Rapoport, A.; Budescu, D.V. Generation of random series in two-person strictly competitive games. *J. Exp. Psychol. Gen.* **1992**, *121*, 352–363. [CrossRef]

30. Forder, L.; Dyson, B.J. Behavioural and neural adaptation of *win-stay* but not *lose-shift* strategies as a function of outcome value. *Sci. Rep.* **2016**, *6*, 33809. [CrossRef]

31. Aczel, B.; Kekees, Z.; Bago, B.; Szollosi, A.; Foldes, A. An empirical analysis of the methodology of automatic imitation research in a strategic context. *J. Exp. Psychol. Hum. Percept. Perform.* **2015**, *41*, 1049–1062. [CrossRef]

32. Mehta, J.; Starmer, C.; Sugden, R. The nature of salience: An experimental investigation of pure coordination games. *Am. Econ. Rev.* **1994**, *84*, 658–673.

33. Kangas, B.D.; Berry, M.S.; Cassidy, R.N.; Dallery, J.; Vaidya, M.; Hackenberg, T.D. Concurrent performance in a three-alternative choice situation: Response allocation in a Rock/Paper/Scissors game. *Behav. Process.* **2009**, *82*, 164–172. [CrossRef]

34. Wang, Z.; Xu, B. Incentive and stability in the Rock-Paper-Scissors game: An experimental investigation. *arXiv* **2014**, arXiv:1407.1170.

35. Stöttinger, E.; Filipowicz, A.; Danckert, J.; Anderson, B. The effects of prior learned strategies on updating an opponent's strategy in the Rock, Paper, Scissors game. *Cogn. Sci.* **2014**, *38*, 1482–1492. [CrossRef]

36. Cournot, A. Recherches sur les principes mathematiques de la theorie des richesse. In *Researches into the Mathematical Principles of the Theory of Wealth*, English ed.; Bacon, N., Ed.; Macmillan: New York, NY, USA, 1897.

37. Lee, D.; Seo, H.; Jung, M.W. Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci.* **2012**, *35*, 287–308. [CrossRef]

38. Thorndike, E.L. *Animal Intelligence*; Macmillan: New York, NY, USA, 1911.

39. Bolles, R.C. Species-specific defense reactions and avoidance learning. *Psychol. Rev.* **1970**, *77*, 32–48. [CrossRef]

40. Stagner, J.P.; Michler, D.M.; Rayburn-Reeves, R.M.; Laude, J.R.; Zentall, T.R. Midsession reversal learning: Why do pigeons anticipate and perseverate? *Learn. Behav.* **2013**, *41*, 54–60. [CrossRef]

41. Sulikowski, D.; Burke, D. Win shifting in nectarivorous birds: Selective inhibition of the learned win-stay responses. *Anim. Behav.* **2012**, *83*, 519–524. [CrossRef]

42. Lyons, J.; Weeks, D.J.; Elliott, D. The gambler's fallacy: A basic inhibitory process? *Front. Psychol.* **2013**, *4*, 72. [CrossRef]

43. Plonsky, O.; Teodorescu, K.; Erev, I. Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychol. Rev.* **2015**, *122*, 621–647. [CrossRef]

44. Soutschek, A.; Schubert, T. The importance of working memory updating in the Prisoner's dilemma. *Psychol. Res.* **2016**, *80*, 172–180. [CrossRef]

45. Hahn, U.; Warren, P.A. Perceptions of randomness: Why three heads are better than four. *Psychol. Rev.* **2009**, *116*, 454–461. [CrossRef]

46. Rayburn-Reeves, R.M.; Laude, J.R.; Zentall, T.R. Pigeons show near-optimal win-stay/lose-shift performance on a simultaneous-discrimination, midsession reversal task with short intertrial intervals. *Behav. Process.* **2013**, *92*, 65–70. [CrossRef]

47. Marshall, A.T.; Kirkpatrick, K. The effects of the previous outcome on probabilistic choice in rats. *J. Exp. Psychol. Anim. Behav. Process.* **2013**, *39*, 24–38. [CrossRef]

48. Elliott, R.; Vollm, B.; Drury, A.; McKie, S.; Richardson, P.; Deakin, J.F.W. Co-operation with another player in a financially rewarded guessing game activates regions implicated in theory of mind. *Soc. Neurosci.* **2006**, *1*, 385–395. [CrossRef]

49. Rayburn-Reeves, R.M.; Molet, M.; Zentall, T.R. Simultaneous discrimination reversal learning in pigeons and humans: Anticipatory and perseverative errors. *Learn. Behav.* **2011**, *39*, 125–137. [CrossRef]

50. Gaissmaier, W.; Schooler, L.J. The smart potential behind probability matching. *Cognition* **2008**, *109*, 416–422. [CrossRef]

51. Tamura, K.; Masuda, N. Win-stay lose-shift strategy in formation changes in football. *EPJ Data Sci.* **2015**, *4*, 9. [CrossRef]

52. Heyes, C.M. Theory of mind in nonhuman primates. *Behav. Brain Sci.* **1988**, *21*, 101–148. [CrossRef]

53. Hachiga, Y.; Schwartz, L.P.; Tripoli, C.; Michaels, S.; Kearns, D.; Silberberg, A. Like chimpanzees (*Pan troglodytes*), pigeons (*Columba livia domestica*) match and nash equilibrate where humans (*Homo sapiens*) do not. *J. Comp. Psychol.* **2018**, *133*, 197–206. [CrossRef]

54. Brauer, J.; Call, J.; Tomasello, M. Chimpanzees really know what others can see in a competitive situation. *Anim. Cogn.* **2007**, *10*, 439–448. [CrossRef]

55. Vlaev, I.; Chater, N. Debiasing context effects in strategic decisions: Playing against a consistent opponent can correct perceptual but not reinforcement biases. *Judgm. Decis. Mak.* **2008**, *3*, 463–475.

56. Dyson, B.J.; Sundvall, J.; Forder, L.; Douglas, S. Failure generates impulsivity only when outcomes cannot be controlled. *J. Exp. Psychol. Hum. Percept. Perform.* **2018**, *44*, 1483–1487. [CrossRef]

57. Weiger, P.; Spaniol, J. The effect of time pressure on risky financial decisions from description and decision from experience. *PLoS ONE* **2015**, *10*, e0123740. [CrossRef] [PubMed]

58. Sanfey, A.G.; Rilling, J.K.; Aronson, J.A.; Nystrom, L.E.; Cohen, J.D. The neural basis of economic decision-making in the Ultimatum game. *Science* **2003**, *300*, 1755–1758. [CrossRef] [PubMed]

59. Van't Wout, M.; Kahn, R.S.; Sanfey, A.G.; Aleman, A. Affective state and decision-making in the Ultimatum Game. *Exp. Brain Res.* **2006**, *169*, 564–568. [CrossRef] [PubMed]

60. Laakasuo, M.; Palomäk, J.; Salmela, M. Emotional and social factors influence poker decision making accuracy. *J. Gambl. Stud.* **2015**, *31*, 933–947. [CrossRef] [PubMed]

61. Palomäki, J.; Laakasuo, M.; Salmela, M. Losing more by losing it: Poker experience, sensitivity to losses and tilting severity. *J. Gambl. Stud.* **2014**, *30*, 187–200. [CrossRef] [PubMed]

62. Mitzenmacher, M.; Upfal, E. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*; Cambridge University Press: Cambridge, UK, 2017.

63. Petry, N.M.; Blanco, C.; Auriacombe, M.; Borges, G.; Bucholz, K.; Crowley, T.J.; Grant, B.F.; Hasin, D.S.; O'Brien, C. An overview of and rationale for changes proposed for pathological gambling in DSM-5. *J. Gambl. Stud.* **2014**, *30*, 493–502. [CrossRef] [PubMed]

64. Clarke, D. Impulsiveness, locus of control, motivation and problem gambling. *J. Gambl. Stud.* **2004**, *20*, 319–345. [CrossRef]

65. James, R.L.; O'Malley, C.; Tunney, R.J. Why are some games more addictive than others: The effects of timing and payoff on perseverance in a slot machine game. *Front. Psychol.* **2016**, *7*, 46. [CrossRef]

66. Larson, M.J.; South, M.; Krauskopf, E.; Clawson, A.; Crowley, M.J. Feedback and reward processing in high-functioning autism. *Psychiatry Res.* **2011**, *187*, 198–203. [CrossRef]

67. McPartland, J.C.; Crowley, M.J.; Perszyk, D.R.; Mukerji, C.E.; Naples, A.J.; Wu, J.; Mayes, L.C. Preserved reward outcome processing in ASD as revealed by event-related potentials. *J. Neurodev. Disord.* **2012**, *4*, 16. [CrossRef]

68. Muller, S.V.; Moller, J.; Rodriguez-Fornells, A.; Munte, T.F. Brain potentials related to self-generated and external information used for performance monitoring. *Clin. Neurophysiol.* **2005**, *116*, 63–74. [CrossRef]

69. Holroyd, C.B.; Hajcak, G.; Larsen, J.T. The good, the bad and the neutral: Electrophysiological responses to feedback stimuli. *Brain Res.* **2006**, *1105*, 93–101. [CrossRef]

70. Gu, R.; Feng, X.; Broster, L.S.; Yuan, L.; Xu, P.; Luo, Y.-J. Valence and magnitude ambiguity in feedback processing. *Brain Behav.* **2017**, *7*, e00672. [CrossRef]

71. Dixon, M.J.; MacLaren, V.; Jarick, M.; Fugelsang, J.A.; Harrigan, K.A. The frustrating effects of just missing the jackpot: Slot machine near-misses trigger large skin conductance responses, but no post-reinforcement pauses. *J. Gambl. Stud.* **2013**, *29*, 661–674. [CrossRef]

72. Ulrich, N.; Hewig, J. Electrophysiological correlates of near outcome and far outcome sequence processing in problem gamblers and controls. *Int. J. Psychophysiol.* **2019**, in press.

73. Miltner, W.H.R.; Brown, C.H.; Coles, M.G.H. Event related brain potentials following incorrect feedback in a time estimation task: Evidence for a generic neural system for error detection. *J. Cogn. Neurosci.* **1997**, *9*, 787–796. [CrossRef]