The Lottery Player's Fallacy

Why Labels Predict Strategic Choices[§]

Irenaeus Wolff

Thurgau Institute of Economics (TWI) / University of Konstanz
Hauptstrasse 90, 8280 Kreuzlingen, Switzerland
wolff@twi-kreuzlingen.ch

Forthcoming in: *Journal of Economic Behavior & Organization*

Abstract:
This paper examines games with non-neutral option labels (such as "A", "B", "A", "A") and finds surprisingly invariant behaviour across games. The behaviour closely resembles the choices people make when they have to bet on one of the options in individual lotteries. An option's 'representativeness' (lack of distinguishing features) and 'reachability' (physical centrality, salience, and valence) determine choice behaviour in both the lotteries and the highly strategic games. There is no evidence of people best-responding to others' betting(-like) behaviour. This is in line with the idea that once people decide that strategic reasoning would not take them any further, they pick an alternative as if they were betting on one of their 'current best-responses'. The findings explain the well-documented seeker advantage in hide-and-seek games, as well as why participants often display behaviour that could be exploited by others. On top, they help understand why in national lotteries, people also tend to bet on identical subsets of the available numbers.

Keywords: Bounded Rationality, Level-k, Salience, Strategic Behaviour, Hide & Seek, Discoordination, Rock-Paper-Scissors, Representativeness.
JEL: C72, C90, D83.

[§]This manuscript supersedes the discontinued project "Lucky Numbers in Simple Games". I am thankful for the encouragement and helpful comments of Colin Camerer, Adrian Chadi, Vincent Crawford, Fabian Dvorak, Sebastian Fehrler, Urs Fischbacher, Susanne Goldlücke, Shaun Hargreaves Heap, Botond Köszegi, Jörg Oechssler, Stefan Penczynski, Ariel Rubinstein, Abdolkarim Sadrieh, Dirk Sliwka, Robert Sugden, Chris Starmer, Marie-Claire Villeval, Roberto Weber, the lively research group at the Thurgau Institute of Economics (TWI), the anonymous referees and (advisory) editors at *JEBO*, and participants of the Game Theory Society's World Congress 2016, the UEA Behavioural Game Theory Workshop 2016, and the ESA World Meeting 2019.

1 Introduction

When people play the lottery, they tend to make similar choices. Over the last three decades, researchers have identified a number of characteristics of the numbers lottery players choose. For example, players tend to choose "situationally available" numbers and numbers in the centre of the lottery ticket, as well as combinations "with an eye for aesthetics" (Wang et al., 2016).[1] Just to point to two extreme examples, on 18th June, 1977, 205 people had to share the jackpot of the German lottery—all betting on the ("situationally available") winning numbers of the Dutch lottery of the week before. Similarly, on 4th October, 1997, 124 people had to share the jackpot, all betting on the same ("aesthetic") U-pattern on their lottery tickets.[2] However, a question that the literature does not seem to address is: Why do people choose these numbers *given that others are choosing them*, too? One potential reason could be that people ignore the fact that they are in a strategic situation: That maximising expected payoffs means selecting not only the right numbers but also numbers that few others choose.

This paper is about situations that are evidently strategic. While game theory describes how perfectly rational players would interact in any strategic situation, its predictive power for actual behaviour should be higher in situations whose strategic nature cannot be ignored. However, it is clear by now that people's actual strategic reasoning differs from the game-theoretic ideal (*e.g.*, Crawford et al., 2013). What remains unclear is what strategic reasoning people do use in important games like pure discoordination games, hide-and-seek, or rock-paper-scissors.

This paper argues that these games (and others) share important features with 'lottery games' like the above, in which players get to choose their numbers and winnings are divided when there are several winners. The evidently strategic situations have a unique symmetric mixed-strategy Nash equilibrium in which players mix uniformly amongst all options. Hence, in equilibrium—much like in the lottery-game context—all options afford the same prospects.[3] Moreover, similarly to the national-lottery data and in contrast to the equilibrium prediction, scholars have documented a bunching of choices (*e.g.*, Rubinstein et al., 1997).[4]

[1]For other recent contributions on how people make lottery-ticket choices, see also, *e.g.*, Kong et al. (forthc.), Krawczyk and Rachubik (2019), or Suetens et al. (2016).

[2]Source: https://www.lottoland.com, last accessed on 11th March, 2019.

[3]Note that the equivalence of the options does not hinge on the particular choice of the strategic-reasoning model. Uniform randomisation is also the quantal-response equilibrium (McKelvey and Palfrey, 1995) given that strategies are not distinguishable by their payoffs. Also, level-k (Nagel, 1995; Stahl and Wilson, 1994) or cognitive-hierarchy models (Camerer et al., 2004) yield the same solution if they rely on a uniformly-mixing level-0.

[4]There is a single model in the literature that accounts for some of Rubinstein et al.'s data. Crawford and Iriberri (2007) show that a level-k variant based on salience as level-0 can account

This paper shows that the bunching of choices follows the same principles as individual-lottery betting and is common even amongst hiders in a hide-and-seek game. However, for hiders choosing the same option as many others—and in particular, the same option as many seekers—is a particularly bad choice. So, what remains to be answered is: Why does behaviour deviate so much from the game-theoretic ideal also in the evidently strategic games, why is it so close to betting behaviour in individual lotteries even when actions are strategic substitutes, and why do so many people leave money on the table by not exploiting others' betting-like behaviour?

Answering these questions is important because it is informative for how people make strategic decisions in their private as well as their professional lifes. This paper focuses on situations that have three characteristics. (i) The situations are one-(or first-)time occasions, so that players cannot condition on any history. (ii) The players' options have non-neutral 'labels', where the term 'label' represents any non-payoff relevant attributes. And (iii) the options are otherwise indistinguishable (which makes the situation a "strategy-isomorphic game", *cf.* Hargreaves Heap et al., 2014).[5]

While it may be unclear whether strategy-isomorphism is a widespread feature in everyday life, it is a feature that is common for many examples in the game-theoretic literature from coordination to hide-and-seek games—and lottery games. With respect to labels, it is clear that in real-life situations, alternatives are hardly ever abstract objects. Instead, alternatives usually have non-neutral labels attached to them, and they often have a spatial ordering. When driving from one city to another, we may choose between the "northern" and the "southern route", but we rarely choose between "option i and option j, where $i \neq j$". It has long been known that such non-neutral frames often have an influence on choices in an individual-choice context, and this has been incorporated into a choice-theoretic framework (*i.a.,* by Salant and Rubinstein, 2008).[6]

for hide-and-seek data from a number of frames. However, Hargreaves Heap et al. (2014) argue that if Crawford and Iriberri (2007) had used also the data from Rubinstein et al.'s coordination and discoordination games, the model's explanatory power would have been small. On the other hand, Penczynski (2016) finds evidence of level-k-like thinking in hide-and-seek games played on the central frame from Crawford and Iriberri (2007), ABAA. At the same time, Wolff (2016) focuses on the same frame and elicits salience in nine different ways. The elicited salience patterns all tend to be similar, but they do not allow to account for the data when used as level-0.

[5]See Alós-Ferrer and Kuzmics (2013) for a treatment of the frame-induced symmetry structure and its consequences for predicted behaviour in such games. Earlier studies on behaviour in this type of environment include Rubinstein et al. (1997, coordination games, discoordination games, and hide-and-seek games), or Mehta et al. (1994) and Bardsley et al. (2010, who restrict their focus to coordination games with options carrying naturally occurring labels, as studied by Schelling, 1960). See also Scharlemann et al. (2001) for trust games where participants' interaction partners were "labeled" by photographs.

[6]Recently, Piccione and Spiegler (2012), Spiegler (2014), and Salant and Siegel (2018) intro-

Finally, many economically important situations are one-time situations or have a first period without precedents (in which case first-period behaviour may be particularly important as in case of multiple equilibria, it may determine the equilibrium that will be played in the long run). As an example, think of competing firms designing a new product who have to adapt one of two distinct but equivalent standards.

While a huge number of experimental studies have combined strategy-isomorphism with one-shot interactions, this paper is one of relatively few that addresses the influence of non-neutral labels as a ubiquitous aspect of reality in strategic situations. To do so, I combine the relatively rich data set from Rubinstein and Tversky (1993) and Rubinstein et al. (1997) with data I gathered for other projects as well as some new data, to arrive at a final data set that includes data from over 2'000 participants.

The answer that this paper offers is what I call the 'lottery player's fallacy'. Once people's (strategic) deliberations bring about the conclusion that there is no reason to favour any of the available options, they enter a 'lottery-player mode'. In the 'lottery-player mode', people simply bet on one of the actions, no longer paying attention to the fact that others may act in a similar way. This implies that such games will yield the same choice pattern as individual lotteries over the same option sets.

Falling prey to the 'lottery player's fallacy' changes the set of best-responses for others and could be exploited by them. Surprisingly, however, I do not find any evidence for participants playing a best-response to others' betting-like behaviour. As this paper shows, a simplistic model incorporating the 'lottery player's fallacy' performs better in out-of-sample predictions (where over-fitting is not an issue; it also is undominated in data fitting) than a whole number of variants that allow for optimal reactions to 'lottery-player-mode choices'.

The 'lottery player's fallacy' provides an explanation also for why coordination-game choices are different from choices in other games. In the coordination games I study, players' strategic deliberations—which may, for example, be based on team-reasoning—will not yield the conclusion that there is no reason to favour any of the options: Choosing salient items yields a higher probability of coordinating. Hence, players will not enter the 'lottery-player mode' in coordination games. And hence, their coordination-game behaviour will differ from behaviour in lotteries as well as the other games, in which people will enter the 'lottery-player mode'.

As a precondition for the offered explanation of the data, this paper estab-

---

duced models incorporating framing also into interactive situations: in these studies, sellers can choose which frame to communicate to their buyers to influence the latters' buying decisions. In the present study, in contrast, players interact under exogenously given sets of labels.

lishes that when people play different games on the same frame, the actual game being played has a surprisingly minor role (where a "frame" refers to the labeling and spatial ordering of the available alternatives). Unless people are playing a coordination game, behaviour always is well-correlated with behaviour in a lottery played on the same frame. This invariance is all the more remarkable given that the participants come from different cultural backgrounds (Israel, US, Germany) and different generations (students in the 1990s vs in the 2010s).

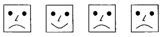
Behaviour in lotteries can be modelled well by the available options' "representativeness" and their "reachability". In this context, representativeness means that the item does not have "prominent, 'non-random' properties" (Bar-Hillel, 2015, citing Teigen, 1983). Reachability includes salience (which makes an option cognitively reachable), valence (which may make an option more attractive to reach out for), and physical proximity (which in our context means centrality: If options are presented on a horizontal line, then the options right in front of the person—the options in the middle—would be easier to reach with one's hands than the options at the extremes).

In Section 4.3, I regress how often an option is chosen in the lottery setting (the option's 'betting attraction') on empirical measures of "representativeness" and on different aspects of "reachability". As predicted by the 'lottery player's fallacy', this regression is indicative of choices in the games, too. In fact, using fitted 'betting attraction' allows to predict strategy choices in games out-of-sample (and partially even 'out-of-subject-pool') virtually as well as measured 'betting attraction'.

2 The data

I use data of several papers on behaviour in games where actions are not distinguishable by their payoffs, complemented by new data from the "to-your-right game" I describe below. All of the games are two-player four-option games played on different frames. The left-hand column of Table 1 provides an overview of the frames. I use all frames presented by Rubinstein and Tversky (1993) and Rubinstein et al. (1997), plus two obvious complements, BAAA and AAAB, as well as the Ace-2-3-Joker frame introduced by O'Neill (1987) and also referred to in Crawford and Iriberri (2007, frames are represented as follows: *E.g.*, BAAA represents a frame in which the left-most option—or "box"—is labelled B and all other options are labelled A).[7]

[7] It can be argued that including all of Rubinstein et al.'s frames distorts the analysis because some of these frames use labels with positive or negative connotations. Therefore, choosing the associated actions may increase or decrease utility on top of the utility associated with the resulting monetary outcome. I nevertheless include all of Rubinstein et al.'s frames, for three reasons:

Frame	Coordination	Discoordination	Hiders	Seekers	To-your-right
	RTH (50)	RTH (49)	RTH (53)	RTH (62)	new (110)
polite-rude-honest- -friendly	RTH (50)	RTH (49)	RTH (53)	RTH (62)	new (110)
	RTH (50)	RTH (49)	RTH (53)	RTH (62)	new (110)
ABAA	RTH+W (122)	RTH+WB +BW (442)	RTH+HW +W (339)	RTH+HW +W (281)	new (110)
	RTH (50)	RTH (49)	RTH (53)	RTH (62)	new (110)
hate-detest-love-dislike	RTH (50)	RTH (49)	RTH (53)	RTH (62)	new (110)
1-2-3-4	RT (184)†	WB+BW (292)	RT (187)	RT (84)	new (110)
AABA	RT (185)†	WB+BW (292)	RT (189)	RT (85)	new (110)
Ace-2-3-Joker		WB+BW (292)			new (110)
BAAA		WB+BW (292)			new (110)
AAAB		WB+BW (292)			new (110)

†Pooled from the "Chooser" and "Guesser" framings. RTH: Rubinstein et al. (1997). RT: Rubinstein and Tversky (1993). HW: Heinrich and Wolff (2012). BW: Bauer and Wolff (2018). WB: Wolff and Bauer (2018). W: Wolff (2015).

Table 1: Origin of the data I use (numbers of observations in parentheses).

Table 1 presents the origin of the data I use in this study, together with the number of observations in parentheses. I added data from a new game ("to-your-right") to the data I had gathered from existing sources (coordination, discoordination, and hide-and-seek games). I added this data in order to test predictive success, to make sure the offered explanation has bite not only on pre-existing data but also on a new data set.[8]

The data for the coordination and hide-and-seek games mostly comes from Rubinstein and Tversky's (1993) and Rubinstein et al.'s (1997) studies, only for the ABAA frame, I have additional observations from other studies for each game. For the discoordination games, Rubinstein et al. (1997) only have data for six of the frames. I complement this with data on a different subset of six frames collected

(i) In my view, understanding behaviour in non-neutral landscapes extends beyond 'neutral non-neutral' landscapes (and it would be difficult to draw the line if we accept the idea that people tend to have lucky numbers); (ii) at least some of the behavioural models described in this paper are meant to apply also under 'truly non-neutral' frames; and (iii) excluding frames with clearly positive or negative connotations in our sample does not change the results meaningfully but leaves us with less statistical power for the analysis.

[8]The payoff matrices of all games are given in Appendix B.

for two studies run by Dominik Bauer and myself. Finally, I collected the data for the to-your-right games specifically for this study. A complete tabulation of all the data can be found in Appendix A. Given the data structure, the setup corresponds to a between-participants design in terms of the player roles/games (the same applies to virtually all additional empirical measures described below). The only within-participants variation is in terms of the frames: many players in the sample played the same game in the same role on different frames.

I ran the to-your-right game as the first part of sessions comprised of three parts, where I described each part only as it started. Only one part was paid. If the first part was payoff relevant, the roll of a die selected one of the to-your-right games for payment. Participants played under all eleven frames with a randomised order, random rematching, and without feedback between games. I described the to-your-right game as follows:

> There are four boxes. You and the other participant choose a box without knowing the decision of the respective other. One of you can obtain a prize of 12 Euros. Who wins depends on the relative position of the two chosen boxes. The participant wins whose box lies to the immediate right of the box of the other participant. If a participant chooses the right-most box, then the other participant wins if he chooses the left-most box. Who does not win obtains a consolation prize of 4 Euros. Of course, it is possible that neither you nor the other participant wins.

Participants had not participated in any other experiments using the same type of non-neutral frames.[9]

I initially set out to describe behaviour using level-k-models based on empirical measures of certain psychologic concepts. For this purpose, I collected data from additional tasks in separate sessions. First, I conducted seven sessions with a total of 140 participants that had a BettingTask as the first of several parts (following the same procedures as with the to-your-right games). In the BettingTask, participants faced the following task:

> In each decision situation you have to choose one out of several boxes. Subsequently, one of the boxes will be randomly selected by the cast of a die. In case the randomly-selected box coincides with the box you chose, you receive 12 Euros. If the two boxes do not coincide, you receive 4 Euros.

Next, I conducted three sessions with a total of 58 participants of a HiderBettingTask. This task differs from the BettingTask only in that participants

[9] I used z-Tree (Fischbacher, 2007) and orsee (Greiner, 2015).

Game/Role	Coordination	Discoordination	Hiders	Seekers	To-your-right
Modes that coincide† (in %)	31.3	50	93.8	93.8	68.2
Spearman correlation of ranks	0.23	0.24	0.63	0.61	0.5
p-value of Spearman test	(0.205)	(0.110)	(< 0.001)	(< 0.001)	(0.001)

Table 2: Relationship of BettingTask data and data from the different player roles. Note: †For the AABA-Frame, the BettingTask data have a split mode on the central A and on B. To account for this, I counted the coincidence with the game data only as 0.5 for this frame.

receive the bigger prize if their choice does not coincide with the randomly-selected box. Finally, I asked 96 participants to rate the options' optic salience (SalienceRating) and 102 participants to rate how well each of the boxes within a frame represented all four boxes within that frame (RepresentRating).[10] In both tasks, participants saw the boxes in the same horizontal line-up as in the other tasks. Below each box, they had a slider (empty at the outset) to indicate the level of optical salience (between "extremely conspicuous", top, and "extremely nondescript", bottom) or representativeness (between "totally representative", top, and "not representative at all", bottom).

Observation 1. Choices in the BettingTask and game data are correlated.

The first row of Table 2 shows that the modal choice in the BettingTask data coincides with far more than the expected 25% under random play for the discoordination game (50%), hiders and seekers (both 94%), to-your-right game (68%) but not so much for the coordination game (31%). To assess the correlation in more detail, I look at the choice patterns, assigning a rank of 1 if a certain option in a specific game under a specific frame was chosen most often, a rank of 2 if that option was chosen second-most often, and so on. This analysis then shows a high Spearman correlation if the game data exhibits a very similar pattern as the BettingTask data under the same frame. The resulting rank correlations are all substantial ($0.23 \leq \rho \leq 0.63$); at the same time, the correlation with the coordination-game data seems the least robust correlation with a p-value of 0.2.[11]

[10] The rating tasks were included in BettingTask and HiderBettingTask sessions similar to a post-experimental questionnaire.

[11] This analysis neglects the limited degrees of freedom of the data: once ranks 1-3 are assigned, the remaining option has to have rank 4. If I use the mean squared rank difference between game data and BettingTask data for each frame, sum the result over all frames, and use that as a test statistic against the expected statistic if ranks were randomly assigned, I obtain even lower p-values: 0.103 for coordinators, 0.064 for discoordinators, and $p < 0.001$ for the remaining roles.

3 Amending standard models to account for the data

Interpreting the evidence from Observation 1, it looks like participants have found a way of playing the coordination game that is not so much driven by the principles that drive their BettingTask choices. At the same time, it looks as if they were relying on those principles in the other games to a large degree. In light of these observations, I put forward the following two-step explanation: In Step 1, players start reasoning strategically in some way that allows them to 'solve' the coordination game, presumably by choosing the most salient option. However, this strategic reasoning does not allow them to single out a unique best-response to what they think their opponent will do in the other games.

Given that after step 1, there are multiple options that seem to yield the same expected payoff—presumably, all four options—in step 2, players would choose amongst these options as amongst "evidently equivalent options" (Bar-Hillel, 2015). As Bar-Hillel argues in her article, choice amongst evidently equivalent options may be understood best in terms of the options' representativeness and their "reachability" (which in our case may be determined by salience, physical centrality, and potentially valence).

At this point, we might argue that once players have reached the end of step 2, they could go on in their strategic reasoning, in a level-k-like fashion: Best-responding to others' choice between evidently equivalent options, best-responding to the best-response, or going even further. The following section explores this possibility in detail, juxtaposing several more strategically sophisticated models to the idea that there simply are two types of players: Uniformly mixing (equilibrium) players and players who choose amongst the evidently equivalent (equilibrium) options by choosing 'BettingNumbers'. Before I present the results, let me briefly describe the models I explore.

As I argued in the introduction, there are no 'standard' alternatives available that would come close to explaining the data. Starting from the observation that the BettingTask choices are surprisingly similar to the data from the different games, I construct three 'conceptually sensible' level-k alternatives based on the BettingTask choices. These level-k alternatives essentially add the possibility that some players best-respond to betting-like behaviour, or to the best-response to betting-like behaviour, or even to higher iterations of best-responding.

In analogy to one of the level-k models, I additionally construct an equilibrium with payoff-perturbations that also builds on the notion that BettingTask choices are indicative of non-monetary utility that different option labels and positions may convey. As benchmarks, I include the standard Nash equilibrium, a standard level-k model with uniformly mixing level-0, and the salience-based level-k of Crawford and Iriberri (2007) with an empirically-defined level-0.

NashEqm. The unique symmetric mixed-strategy equilibrium has both players randomise uniformly over all locations. Hence, the likelihood function is:

$$L = \Pi_{i \subset \{H,S,C,D,T\}} \Pi_{f=1,\dots,F_i} \Pi_{j=1,2,3,4} \left(\frac{1}{4}\right)^{X_{fj}^{(i)}},$$

where H, S, C, D, and T stand for hiders, seekers, coordinators, discoordinators, and to-your-right players, respectively, f is the frame, j is the location within frame f, and $X_{fj}^{(i)}$ is the number of players choosing j in role i under frame f.

LuckyNoEqm. A Nash-equilibrium variant in which participants derive extra utility from choosing certain locations (cf. Crawford and Iriberri, 2007). For this model, I interpret the BettingTask data as a measure of participants' inherent preferences for the different locations.[12] I compute utility values from the BettingTask data and re-define the game in terms of these utility values: A multinomial-logit utility model estimated by maximum likelihood yields utility values that I transform in an affine-linear way (to obtain positive utility values). Concretely, relating the latent utility of betting on location j in frame f to the probability of choosing it in the betting task by a logit-choice function yields

$$Pr(j, f) = \frac{e^{\lambda U_b(f,j)}}{\sum_{l=1,2,3,4} e^{\lambda U_b(f,l)}}, \tag{1}$$

where $U_b(f, j)$ is the latent utility value of betting on location j in frame f. The corresponding likelihood function is

$$L = \Pi_{f=1,\dots,F_b} \Pi_{j=1,2,3,4} \frac{e^{\lambda U_b(f,j) X_{fj}^{(b)}}}{(\sum_{l=1,2,3,4} e^{\lambda U_b(f,l)})^{X_{fj}^{(b)}}},$$

where $X_{fj}^{(b)}$ is the number of participants choosing location j in frame f of the betting task. A maximum-likelihood estimation yields estimates for $\lambda U_b(f, j)$ that sometimes happen to be negative. However, I can transform the estimates for $U_b(f, j)$ into $U_b'(f, j) = aU_b(f, j) + b$ so that all $U_b'(f, j)$ are positive.

The predicted choices for the betting task remain the same as in equation (1).[13] However, the transformation does affect the mixed-strategy equilibria for

[12] Of course, this assumes that people are homogeneous in what utilities they derive from the different locations. This is a strong assumption, but it is the best approximation that I have.

[13] To see that the predicted choices remain the same, note that $\lambda U_b(f, j) = \lambda' U_b'(f, j) - \frac{\lambda b}{a}$ for $\lambda' = \frac{\lambda}{a}$. Factoring out $e^{-\lambda b/a}$ in both the numerator and the denominator yields the equivalence.

the games that result when the non-zero entries in the standard game matrix are replaced by the transformed utility values. Assuming that participants will play the mixed equilibrium strategy with probability $(1-\varepsilon)$ and choose a random action with probability ε—not all equilibrium strategies have full support—leads to the following likelihood function:

$$L = \Pi_{i \subset \{H,S,C,D,T\}} \Pi_{f=1,\ldots,F_i} \Pi_{j=1,2,3,4} \left((1-\varepsilon)\pi^*_{ifj}(a,b) + \frac{\varepsilon}{4} \right)^{X^{(i)}_{fj}}, \qquad (2)$$

where $\pi^*_{ifj}(a,b)$ is the equilibrium probability of choosing location j under frame f in game i for the betting-utility transformation given by a and b. The log of (2) is maximised over ε, a, and b.

Take the example of a discoordination game played on ABAA: (Absolute) choice frequencies in the BettingTask were 18, 46, 58, and 18. If these frequencies are the result of a multinomial-logit choice process, the maximum-likelihood estimates for utilities are -0.53, 0.41, 0.64, and -0.53. Using, for example, $a = \frac{2}{3}$ and $b = 1$ those utilities are recalibrated to 0.64, 1.27, 1.43, and 0.64 (which are still in accordance with the BettingTask choice frequencies). We can then use the recalibrated values as the corresponding entries in the normal form game: When a participant chooses one of the end-As and her opponent chooses another location, the participant's utility will be 0.64. Likewise, when she successfully discoordinates by choosing B, her utility will be 1.27. Using the resulting normal-form game, the unique symmetric-equilibrium (mixed) strategy would be (0, 0.47, 0.53, 0). As I point out above, I allow for errors and allow the maximum-likelihood procedure to optimise over the utility-recalibration for the model comparison in part 4.1.

Standard Lk. The standard level-k model with a uniformly mixing level-0, together with the standard auxiliary assumption that the choices of each higher level also leads to a uniform distribution in the aggregate. The predictions of the model (and thus, also the likelihood function) coincide with those of the standard Nash equilibrium for the games examined here. Therefore, the model will be subsumed under "NashEqm" for the remainder of the paper.

Salience-Lk. Crawford and Iriberri's (2007) level-k model in which level-0 follows salience, and level-k players with $k > 0$ play a best-response to level-$(k-1)$ players. Rather than making assumptions about what is salient, I use data from the SalienceRating task as the level-0 to base the model on.[14] The

[14] I use the distribution of locations that participants ranked as most salient, to obtain a metric that is comparable to the data from the BettingTask. Using the average salience rating for each location does not change the results in any significant way.

corresponding likelihood function is the following:

$$L = \Pi_{i\subset\{H,S,C,D,T\}}\Pi_{f=1,\ldots,F_i}\Pi_{j=1,2,3,4}\left[r\frac{X^{\text{SAL}}_{fj}}{\sum_m X^{\text{SAL}}_{fm}}+s\rho^{\text{SAL},1}_{ifj}+t\rho^{\text{SAL},2}_{ifj}+u\rho^{\text{SAL},3}_{ifj}+v\rho^{\text{SAL},4}_{ifj}\right]^{X^{(i)}_{fj}},$$

where X^{SAL}_{fj} is the number of SalienceRating respondents for whom location j was the most salient location under frame j, and $\rho^{\text{SAL},k}_{ifj}$ is the probability that a player of level k chooses location j under frame f in player role i.

For our discoordination-game example on ABAA, the first A is held to be the most salient location by 2% of all SalienceRating participants, B by 91%, and the central and last As by 4% each. Therefore, we would expect level-0 to choose with probabilities $(0.02, 0.91, 0.04, 0.04)$, uneven levels to choose the first A, and even levels to randomise between the other three locations ($\rho^{\text{SAL},1}_{ifj}$ and $\rho^{\text{SAL},3}_{ifj}$ coincide for the discoordination game, as do $\rho^{\text{SAL},2}_{ifj}$ and $\rho^{\text{SAL},4}_{ifj}$).

Betting-Lk. This level-k model uses as level-0 the data from the BettingTask. In level-k theories, level-0 is supposed to be people's intuitive reaction to the game, which may well coincide with the choice they make in a lottery. The likelihood function is given by:

$$L = \Pi_{i\subset\{H,S,C,D,T\}}\Pi_{f=1,\ldots,F_i}\Pi_{j=1,2,3,4}\left[r\frac{X^{\text{BET}}_{fj}}{\sum_m X^{\text{BET}}_{fm}}+s\rho^{\text{BET},1}_{ifj}+t\rho^{\text{BET},2}_{ifj}+u\rho^{\text{BET},3}_{ifj}+v\rho^{\text{BET},4}_{ifj}\right]^{X^{(i)}_{fj}},$$

where X^{BET}_{fj} is the number of respondents choosing location j in frame j of the BettingTask.

In the ABAA-discoordination-game example, betting proportions—and hence, level-0 choices—are 13%, 33%, 41%, and 13%. Then, uneven levels randomise between the end As and even levels between the two locations in the middle.

asymmBetting-Lk. This parallels Crawford and Iriberri' 2007 level-k model with an asymmetric level-0 that follows 'betting attraction' for coordinators and seekers. For discoordinators, hiders, and to-your-right players, it instead follows an inverted version of the betting proportions ('avoids betting attraction'), leading to the following likelihood function:[15]

$$L = \Pi_{i\subset\{S,C\}}\Pi_{f=1,\ldots,F_i}\Pi_{j=1,2,3,4}\left[r\frac{X^{\text{BET}}_{fj}}{\sum_m X^{\text{BET}}_{fm}} + s\rho^{\text{BET},1}_{ifj} + t\rho^{\text{BET},2}_{ifj} + u\rho^{\text{BET},3}_{ifj} + v\rho^{\text{BET},4}_{ifj}\right]^{X^{(i)}_{fj}} + \\ + \Pi_{i\subset\{H,D,T\}}\Pi_{f=1,\ldots,F_i}\Pi_{j=1,2,3,4}\left[rx^{\text{NBET},0}_{fj} + s\rho^{\text{NBET},1}_{ifj} + t\rho^{\text{NBET},2}_{ifj} + u\rho^{\text{NBET},3}_{ifj} + v\rho^{\text{NBET},4}_{ifj}\right]^{X^{(i)}_{fj}},$$

where $x^{\text{NBET},0}_{fj} = (1 - \frac{X^{\text{BET}}_{fj}}{\sum_m X^{\text{BET}}_{fm}})/3$ and $\rho^{\text{NBET},2}_{ifj}$ are the probabilities that follow from using $x^{\text{NBET},0}_{fj}$ as level-0. In the ABAA-discoordination-game example, level-0 choices therefore equal 29%, 22%, 20%, and 29%. Then, uneven levels choose the central A and even levels randomise between the other three locations.

Bounded Lk. This model differs from standard level-k with a uniformly randomising level-0 only in terms of level-1. It incorporates that level-1 players may respond to uniform randomisation by non-uniform randomisation (or by not randomising at all): The BettingTask and HiderBettingTask elicit what participants do when facing uniform randomisation. Level-1 seekers will act like participants in the BettingTask, whereas level-1 discoordinators and level-1 hiders will act like participants in the HiderBettingTask. For the to-your-right games, I also use the HiderBettingTask given that participants have to discoordinate with a random choice also in this game, even if in a very specific way. The resulting likelihood function is:

$$L = \Pi_{i\subset\{S,C\}}\Pi_{f=1,\ldots,F_i}\Pi_{j=1,2,3,4}\left[\frac{r}{4} + s\frac{X^{\text{BET}}_{fj}}{\sum_m X^{\text{BET}}_{fm}} + t\rho^{\text{BD},2}_{ifj} + u\rho^{\text{BD},3}_{ifj} + v\rho^{\text{BD},4}_{ifj}\right]^{X^{(i)}_{fj}} + \\ + \Pi_{i\subset\{H,D,T\}}\Pi_{f=1,\ldots,F_i}\Pi_{j=1,2,3,4}\left[\frac{r}{4} + s\frac{X^{\text{HBET}}_{fj}}{\sum_m X^{\text{HBET}}_{fm}} + t\rho^{\text{BD},2}_{ifj} + u\rho^{\text{BD},3}_{ifj} + v\rho^{\text{BD},4}_{ifj}\right]^{X^{(i)}_{fj}},$$

where X^{HBET}_{fj} is the number of participants choosing location j in frame j of the HiderBettingTask.

For ABAA, the HiderBettingTask choice frequencies are 9%, 53%, 21%, and 17%. Therefore, in our discoordination-game example, level-0 would randomise uniformly, level-1 would choose with probabilities $(0.09, 0.53, 0.21, 0.17)$, even levels would choose the first A, and uneven levels of level-3 or higher would randomise uniformly among all locations but the first A.

[15]Grouping the to-your-right players with discoordinators and hiders rests on the idea that to-your-right is a specific type of discoordination game as well. If I was to group the to-your-right players with coordinators and seekers instead, the model would be much worse at fitting the to-your-right data, and gain ever so slightly in terms of the out-of-game predictions in Section 4.2 (with a mean squared prediction error of 0.0097 instead of 0.0100).

model	fitted on	LogL	MSE	parameters†
Standard Lk	coordination	-1027	0.0782	—
(asymm)Betting-Lk		-951	0.0616	2 $(r, s+t+u+v)$
Bounded Lk		-945	0.0635	3 $(\varepsilon, s, t+u+v)$
Salience-Lk		-885	0.0154	2 $(r, s+t+u+v)$
Betting-Lk	discoordination	-2980	0.0056	3 $(r, s+u, t+v)$
NashEqm/Standard Lk		-2975	0.0045	—
asymmBetting-Lk		-2974	0.0052	5 (r, s, t, u, v)
Salience-Lk		-2972	0.0039	3 $(r, s+u, t+v)$
Bounded Lk		-2967	0.0034	3 $(\varepsilon, s+u, t+v)$
LuckyNoEqm		-2960	0.0038	3 (a, b, ε)
BettingNumbers + Uniform		-2959	0.0039	1 (ε)
NashEqm/Standard Lk	hide & seek	-2412	0.0160	—
Salience-Lk		-2404	0.0126	5 (r, s, t, u, v)
LuckyNoEqm		-2361	0.0116	3 (a, b, ε)
Bounded Lk		-2340	0.0089	5 $(\varepsilon, s, t, u, v)$
BettingNumbers + Uniform		-2316	0.0076	1 (ε)
asymmBetting-Lk		-2300	0.0064	5 (r, s, t, u, v)
Betting-Lk		-2297	0.0066	5 (r, s, t, u, v)
NashEqm/Standard Lk	to-your-right	-1677	0.0034	—
Bounded Lk		-1676	0.0033	5 $(\varepsilon, s, t, u, v)$
BettingNumbers + Uniform		-1674	0.0030	1 (ε)
Betting-Lk		-1672	0.0028	5 (r, s, t, u, v)
Salience-Lk		-1672	0.0028	5 (r, s, t, u, v)
LuckyNoEqm		-1669	0.0039	3 (a, b, ε)
asymmBetting-Lk		-1665	0.0021	5 (r, s, t, u, v)

†In Level-k models, r-v represent the proportions of levels 0-4; a and b are the parameters used for the affine-linear transformation of utilities in the equilibrium with payoff perturbation; and ε stands for any mixture with uniform randomisation (due to errors or otherwise).

Table 3: Performance of the models in terms of data fitting, ordered by log-likelihood.

4 Results

4.1 Accounting for behaviour

The coordination-game data is explained best by participants choosing what they see as the most salient option, as evidenced by the success of the Salience-Lk model in the top part of Table 3 (Salience-Lk has the largest log-likelihood and the lowest mean squared error; I omit the Nash-equilibrium model here as it does not make a unique prediction). For virtually all frames we use, this would be in line with team reasoning (following the decision rule that, when followed by all players, yields the best outcome) as much as with the SALIENCE-Lk model.[16] Hence, the data would be consistent also with the two-step reasoning I propose in this paper if we assume that players use team reasoning as step 1.

The focus of this paper, however, will be on the second step of the proposed reasoning and on the question of whether including a best-response assumption to the introduction of 'choice as in a BETTINGTASK' will allow to account for the data better. For the set of label sets used in this paper, Section 4.3 shows that BETTINGTASK choices are determined by 'representativeness' (the lack of distinguishing features) and 'reachability' (physical centrality, salience, and valence), in line with the psychology literature.

Result 1. Adding 'BETTINGNUMBERS' to a model generally improves that model's ability to fit the data. However, there is no clear evidence that additionally allowing for best-responses to 'BETTINGNUMBER'-choosing players would yield any further improvement.

Looking at Table 3, the amended models (BETTING-Lk, BOUNDED Lk, and LUCKYNOEQM) generally outperform their corresponding standard model (STANDARD Lk and NASHEQM) both in terms of exhibiting a larger log-likelihood and a smaller mean squared error (MSE). The single exceptions is that in the discoordination game, the (ASYMM)BETTING-Lk models perform worse than STANDARD Lk.

Looking at the discoordination game (second part of Table 3), the BETTINGNUMBERS + UNIFORM mix exhibits the largest log-likelihood of the models. It is outperformed in terms of the mean squared error (MSE) by SALIENCE-Lk, BOUNDED Lk, and LUCKYNOEQM. However, each of these models performs clearly worse than BETTINGNUMBERS + UNIFORM when fitted on hide-and-seek data, as the third part of Table 3 shows. Here, (ASYMM)BETTING-Lk—which BETTINGNUMBERS + UNIFORM outperformed clearly in fitting the discoordination data—take on the role of the main contender, with higher log-likelihood and lower MSE.[17]

In the to-your-right games, finally, the ASYMMBETTING-Lk performs best (yielding a distribution of reasoning levels that is almost close to uniform, with 21%, 29%, 29%, and 22% of levels 1 through 4, respectively). The other models tend to be very close to each other in terms of their ability to fit the data. The LUCKYNOEQM model performs second best in terms of the log-likelihood (but worst in

[16]For a formal treatment of team reasoning, I refer the interested reader to Sugden (1995). Roughly, team reasoning requires a choice between decision rules. In the model of Sugden (1995), these decision rules are constructed in a hypothetical state before the labels are assigned to strategies. Decision rules could be "choose the smallest number", "choose your favourite colour", "choose the item standing out the most", or "choose the most representative item". The predicted decision rule should then be unique in being collectively optimal.

[17]Note, however, that the fitted ASYMMBETTING-Lk shows a completely implausible W-pattern in the estimated level-distribution, with 34%, 6%, 28%, 6%, and 25% for levels 0 through 4.

terms of the MSE), while BettingNumber + Uniform performs slightly worse than Betting-Lk and Salience-Lk.

So, while BettingNumbers + Uniform does not dominate any of the other models, it beats each model on at least one data set. It fits the discoordination game best, it fits the hide-and-seek data third best, and in the to-your-right game, it fits the data almost as well as Betting-Lk and Salience-Lk (the comparison to the LuckyNoEqm model is ambiguous because of that model's large MSE).[18]

On top, BettingNumbers + Uniform has only one free parameter as opposed to three (LuckyNoEqm + the amended level-k models for the discoordination data) or five (the amended level-k models for the hide-and-seek and to-your-right data). As further suggestive evidence, the fitted Bounded Lk has virtually only levels 0 (uniform randomisation) and 1 (BettingTask/HiderBettingTask), no matter which game the model is fitted on (a combined 100% if fitted on discoordination, 88% if fitted on hide-and-seek, 94% if fitted on to-your-right). When models have multiple free parameters, a good fit to the data may be due to two different reasons: high explanatory power or over-fitting. To tell apart high explanatory power and over-fitting, I look at out-of-sample predictions in the next section.

4.2 Predicting out of sample

Table 4 displays the weighted mean squared prediction errors for the different models for a leave-one-out procedure. Under this procedure, I fit each model to all games but one and calculate the mean squared prediction error for the left-out game. I predict all games' behaviour in this way, and then weight the resulting mean squared prediction errors by the number of frames for which I have the data. I leave out the coordination games, as for some models the prediction is unclear, and as my focus lies on the games that are *not* coordination games.

Result 2. Adding best-responding types to a mix of BettingNumbers and uniform mixing *worsens* the model's predictive power.

Table 4 shows that BettingNumbers + Uniform predicts out of sample best.[19] In contrast, Bounded Lk and Betting-Lk—which are highly similar to a combi-

[18]The picture (obviously) does not change if we also allow for role-dependent level distributions in a Betting-Lk or a asymmBetting-Lk model. For the symmetric games, the latter models are so-to-speak role-dependent already. And in the hide-and-seek game, the LogL is -2283/-2300 instead of -2297 but the MSE increases substantially (0.0140/0.0088 instead of 0.0066). On top, for the model with a symmetric level-0, I again obtain a level distribution for hiders that resembles the W-pattern I found for asymmBetting-Lk, while the model with an asymmetric level-0 yields a W-pattern for both roles (though far less pronounced than in the other models).

[19]The margin increases further when I treat hiders and seekers as separate data sets (which puts more weight on the hide-and-seek game due to the weighting of the average). In that case,

nation of BettingNumbers + Uniform with additional layers of best-response behaviour—are outperformed by the standard NashEqm/ Standard Lk models. This means that either the prevalence of best-response behaviour is very different between games—or even negligible altogether.

Model	Mean Squared Prediction Error
Salience-Lk	0.0332
asymmBetting-Lk	0.0100
Betting-Lk	0.0085
LuckyNoEqm	0.0082
Bounded Lk	0.0074
NashEqm/Standard Lk	0.0071
BettingNumbers + Uniform	0.0065
predicted BettingNumbers + Uniform	0.0066

Table 4: Mean squared prediction errors from a leave-one-out procedure: I fit the parameters of each model on all games but the one for which the prediction errors are calculated. I do this so that each game's behaviour is predicted once. I report the average (weighted by the number of frames) squared prediction errors over all predictions. The predicted BettingNumbers + Uniform uses the options' fitted 'betting attraction' using the regression from Section 4.3 instead of the BettingNumbers.

4.3 'BettingNumbers': Representativeness & reachability

When motivating why players should be using 'BettingNumbers' in strategy-isomorphic games, I have argued that players use 'BettingNumbers' when they face "evidently equivalent" best-responses to their opponents' candidate strategies. Further, I have argued that we can find an inspiration in psychology research as to what may be influencing choice amongst "evidently equivalent items" (Bar-Hillel, 2015). In this section, I relate 'BettingNumbers' to the psychologic concepts that have been suggested to determine choices in BettingTask-like setups, and that by the 'lottery player's fallacy' hypothesis should also be determining choice in the games.

From her review of the empirical literature, Bar-Hillel (2015) concludes that choice amongst "evidently equivalent items" is likely to be governed by the items'

BettingNumbers + Uniform exhibits a mean squared prediction error of 0.0068, followed by BoundedLk with 0.0082. NashEqm/Standard Lk, predicting particularly badly in hide-and-seek, then comes in fourth with 0.0090.

	coefficient	s.e.	p-value
(Intercept)	-0.45	(0.08)	$5 \cdot 10^{-6}$
negative	0.03	(0.12)	0.8251
positive	-0.05	(0.04)	0.2361
average SalienceRating	0.25	(0.07)	0.0010
average RepresentRating	0.16	(0.07)	0.0404
relative position	0.46	(0.07)	$1 \cdot 10^{-7}$
(relative position)^2	-0.39	(0.06)	$2 \cdot 10^{-7}$
negative · average SalienceRating	-0.24	(0.23)	0.3110
positive · average SalienceRating	0.16	(0.07)	0.0199
R^2	0.78		
AIC	-97.91		
Num. obs.	44 (11 frames with 4 options each)		

Table 5: Regression of relative choice frequencies in the BettingTask (the options' 'betting attraction') on label characteristics (generalised linear model with a power-link function where $\lambda = 0.2$).

representativeness (measured in RepresentRating) and their "reachability". Reachability includes salience (measured in SalienceRating), valence (positive, negative, or neutral), and physical centrality (measured by using an option's relative position—0.5 for the middle, 1 for the right-most location—together with its squared term). Table 5 reports an exploratory regression using the above variables. The regression uses a generalised linear model with a power-link function where $\lambda = 0.2$ was the estimated Box-Cox transformation parameter. While this model does not account for the fact that we are dealing with distributions, it allows for a straightforward interpretation of the coefficients, at the same time accounting for the heteroskedasticity in the original BettingTask data.

Result 3. Representativeness and "Reachability" (valence, salience, and physical centrality) account for what attracts choices in a BettingTask.

As we can see from Table 5, the four psychological concepts that Bar-Hillel (2015) hypothesised should play a role in choice amongst "evidently equivalent items" do indeed account for what options participants choose in the BettingTask: (i) Being rated as more salient ("average SalienceRanking"), (ii) being positioned in the middle (to see this, combine "relative position" and its square); (iii) having a positive connotation when the item is salient ("positive · average SalienceRating"; in the context of our frames, this essentially means that the positively connoted item is presented along with three negatively connoted items); and (iv) being rated as being representative ("average RepresentRating").

While this analysis has to be taken with caution for a whole number of reasons (*i.a.*, because the set of frames is small and rather peculiar), it is in accordance with what seems to determine choices in lottery games in general. Furthermore, when I use the fitted values instead of the empirical 'BettingNumbers' in the out-of-sample prediction from Section 4.2, the mean squared prediction error is still clearly lower than that of any of the level-k variants (see the final row in Table 4). The fact that the predictive power hardly suffers at all supports the idea that the regression does pick up what is crucial about the frames.

5 Discussion

In this paper, I have looked at games in which players' strategies are indistinguishable once we remove the strategies' labels. Many of these games are at the heart of game-theoretic conceptualisations of the world we live in, capturing important elements of every-day life. This paper focuses on versions of the games that incorporate a key aspect of reality, namely that options carry *descriptive* labels. As I have pointed out, the literature currently offers no satisfying explanation for behaviour in these games, in particular when it comes to games that are not pure coordination games.

This paper starts out by looking at the data from several studies on a number of such "strategy-isomorphic games" with non-neutral labels. My first observation is that with the exception of coordination games, behaviour in these games is correlated clearly. Furthermore, the behaviour is correlated clearly with behaviour in individual lotteries played on the same frames. Following psychological research on choice among "evidently equivalent items" (Bar-Hillel, 2015), I regress the lottery choices (the items' 'betting attraction') on measures of the options' "representativeness" (the absence of prominent, 'non-random' properties) and their "reachability" (salience, valence, and physical centrality).

It turns out that the options' (fitted) 'betting attraction' is a strong predictor of behaviour also in the games. This is surprising because it is an exploitable pattern. All the more surprising, I do not find any clear evidence of people best-responding to the behavioural pattern. None of the models that allow for types that best-respond to 'choice by betting attraction' comes even close to matching the out-of-sample predictive power of a simple mixture of uniformly mixing players (as in equilibrium or in standard level-k) with players choosing by 'betting attraction'. In fact, the closest level-k contestant in out-of-sample predictions places virtually no weight on any of the best-responding levels.

These results are consistent with the idea that a substantial fraction of people concludes after an initial strategic-reasoning phase that any option is as good as any other, and that therefore, strict strategic reasoning is no longer needed. Instead, these people choose to 'take a chance and bet on' one of the—now evidently equivalent—options. This simplistic decision strategy accounts for both facts: That (i) people follow exploitable patterns in experimental implementations of strategy-isomorphic games with non-neutral frames and (ii) that these exploitable patterns are correlated accross games.

The facts imply the well-documented seeker advantage in hide-and-seek games (*e.g.*, Rubinstein et al., 1997, or Eliaz and Rubinstein, 2011): If both hiders and seekers display a mixture of uniform mixing and choices that in the end are determined by representativeness and reachability, seekers will win more often than under the uniform-mixture Nash equilibrium. From the literature, we know that the seeker advantage can be reversed when the prize is a bad (when the hider hides a "mine" that a "seeker" should avoid, Rubinstein et al., 1997, or Penczynski, 2016). This observation would follow from the proposed decision strategy, too. Apart from this, the evidence on the reasons for the seeker advantage is ambiguous: being a second-mover seems to lead to more sophisticated reasoning (Penczynski, 2016).[20] However, being a second-mover by itself does not lead to a payoff advantage (Eliaz and Rubinstein, 2011, this latter finding also would follow from the proposed decision strategy).

Finally, if players virtually stop thinking strategically about their opponents even in evidently strategic situations, they are all the more likely to do so in a national-lottery context: The strategic element of the lottery situation seems much less obvious compared to a hide-and-seek game, for example. Thus, the sketched decision strategy not only gives a reason for why representativeness and reachability determine choice in what should be highly strategic situations. The 'lottery player's fallacy' also provides an explanation for two long-standing empirical puzzles—the seeker advantage in hide-and-seek games and the fact that players in national lotteries do not shy away from choosing the same numbers as others.

[20] This is not in line with my estimates for a BETTING-Lk model with role-dependent level distributions but symmetric level-0 (cf. footnote 18), while a fitted model with role-dependent distributions and asymmetric level-0 shows a tiny difference in the above direction when measured by average levels (there is no difference in the medians; *ibd.*).

References

Alós-Ferrer, C. and Kuzmics, C. (2013). Hidden symmetries and focal points. *Journal of Economic Theory*, 148(1):226–258.

Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science*, 10(4):419–433.

Bardsley, N., Mehta, J., Starmer, C., and Sugden, R. (2010). Explaining focal points: Cognitive hierarchy theory *Versus* team reasoning. *Economic Journal*, 120:40–79.

Bauer, D. and Wolff, I. (2018). Biases in beliefs: Experiment evidence. TWI Research Paper Series, No. 109.

Camerer, C. F., Ho, T.-H., and Chong, J.-K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics*, 119(3):861–898.

Crawford, V. P., Costa-Gomes, M. A., and Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51(1):5–62.

Crawford, V. P. and Iriberri, N. (2007). Fatal attraction: Salience, naïveté, and sophistication in experimental 'hide-and-seek' games. *American Economic Review*, 97(5):1731–1750.

Eliaz, K. and Rubinstein, A. (2011). Edgar allan poe's riddle: Framing effects in repeated matching pennies games. *Games and Economic Behavior*, 71(1):88–99.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Greiner, B. (2015). An online recruitment system for economic experiments. *Journal of the Economic Science Association*, 1(1):114–125.

Hargreaves Heap, S., Rojo Arjona, D., and Sugden, R. (2014). How portable is level-0 behavior? A test of level-k theory in games with non-neutral frames. *Econometrica*, 82(3):1133–1151.

Heinrich, T. and Wolff, I. (2012). Strategic reasoning in hide-and-seek games: A note. Research Paper 74, Thurgau Institute of Economics.

Kong, Q., Granić, G. D., Lambert, N. S., and Teo, C.-P. (forthc.). Judgment error in lottery play: When the hot-hand meets the gambler's fallacy. *Management Science*.

Krawczyk, M. W. and Rachubik, J. (2019). The representativeness heuristic and the choice of lottery tickets: A field experiment. *Judgment and Decision Making*, 14(1):51–57.

McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6 – 38.

Mehta, J., Starmer, C., and Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *American Economic Review*, 84(3):658.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences of the United States of America*, 84:2106–2109.

Penczynski, S. P. (2016). Strategic thinking: The influence of the game. *Journal of Economic Behavior & Organization*, 128:72–84.

Piccione, M. and Spiegler, R. (2012). Price competition under limited comparability. *Quarterly Journal of Economics*, 127(1):97–135.

Rubinstein, A. and Tversky, A. (1993). Naive strategies in zero-sum games. Working paper 17-93, The Sackler Institute of Economic Studies.

Rubinstein, A., Tversky, A., and Heller, D. (1997). Naive strategies in competitive games. In Albers, W., Güth, W., Hammerstein, P., Moldovanu, B., and van Damme, E., editors, *Understanding Strategic Interaction–Essays in Honor of Reinhard Selten*, pages 394–402. Springer-Verlag.

Salant, Y. and Rubinstein, A. (2008). (a, f): Choice with frames. *The Review of Economic Studies*, 75(4):1287–1296.

Salant, Y. and Siegel, R. (2018). Contracts with framing. *American Economic Journal: Microeconomics*, 10(3):315–346.

Scharlemann, J., Eckel, C., Kacelnik, A., and Wilson, R. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, 22(5):617–40.

Schelling, T. C. (1960). *The Strategy of Conflict.* Harvard University Press, Cambridge, Massachusetts.

Spiegler, R. (2014). Competitive framing. *American Economic Journal: Microeconomics*, 6(3):35–58.

Stahl, D. O. and Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization*, 25(3):309–327.

Suetens, S., Galbo-Jørgensen, C. B., and Tyran, J.-R. (2016). Predicting lotto numbers: A natural experiment on the gambler's fallacy and the hot-hand fallacy. *Journal of the European Economic Association*, 14(3):584–607.

Sugden, R. (1995). A theory of focal points. *Economic Journal*, 105(430):533–50.

Teigen, K. H. (1983). Studies in subjective probability i: Prediction of random events. *Scandinavian Journal of Psychology*, 24:13–25.

Wang, T. V., Potter van Loon, R. J. D., van den Assem, M. J., and van Dolder, D. (2016). Number preferences in lotteries. *Judgment and Decision Making*, 11(3):243–259.

Wolff, I. (2015). Foundations of strategic thinking and strategic behaviour. Unpublished.

Wolff, I. (2016). Elicited salience and salience-based level-k. *Economics Letters*, 141:134–137.

Wolff, I. and Bauer, D. (2018). Elusive beliefs: Why uncertainty leads to stochastic choice and errors. TWI Research Paper Series, No. 111.

Appendix A Full data

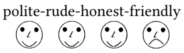
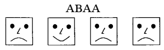
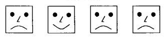
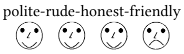
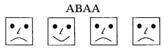
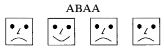
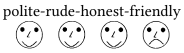
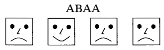
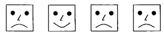
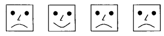
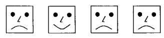
Player role	frame	location 1	location 2	location 3	location 4
coordinators		86	0	10	4
	polite-rude-honest-friendly	6	54	12	28
		6	6	14	74
	ABAA	14	72	13	1
		6	88	6	0
	hate-detest-love-dislike	2	6	88	4
	1-2-3-4	38	17	29	15
	AABA	5	27	54	14
discoordinators		39	14	18	29
	polite-rude-honest-friendly	28	20	32	20
		17	27	23	33
	ABAA	18	21	38	24
		17	40	29	15
	hate-detest-love-dislike	16	29	26	29
	1-2-3-4	21	32	30	17
	AABA	26	24	32	18
	Ace-2-3-Joker	31	17	21	31
	BAAA	34	23	19	23
	AAAB	31	22	18	29
hiders		23	23	43	11
	polite-rude-honest-friendly	15	26	51	8
		21	26	34	19
	ABAA	15	29	33	23
		15	40	34	11
	hate-detest-love-dislike	11	23	38	28
	1-2-3-4	25	22	36	18
	AABA	22	35	19	25
seekers		29	24	42	5
	polite-rude-honest-friendly	8	40	40	11
		7	25	34	34
	ABAA	9	21	53	17
		16	55	21	8
	hate-detest-love-dislike	20	21	55	14
	1-2-3-4	20	18	48	14
	AABA	13	51	21	15
to-your-right players		15	30	32	24
	polite-rude-honest-friendly	22	22	33	24
		18	22	33	27
	ABAA	15	19	34	32
		16	20	33	31
	hate-detest-love-dislike	23	17	30	30
	1-2-3-4	17	21	39	23
	AABA	22	23	29	26
	Ace-2-3-Joker	25	23	31	21
	BAAA	15	26	34	25
	AAAB	24	23	28	25

Table A.1: Full data of the games (relative choice frequencies).

Task	frame	location 1	location 2	location 3	location 4
BettingTask		25	29	36	11
	polite-rude-honest-friendly	12	8	53	27
		16	34	35	16
	ABAA	13	33	41	13
		6	58	26	11
	hate-detest-love-dislike	7	12	69	12
	1-2-3-4	18	21	36	25
	AABA	13	36	36	15
	Ace-2-3-Joker	33	14	19	34
	BAAA	27	25	24	24
	AAAB	18	21	34	26
HiderBettingTask		40	19	19	22
	polite-rude-honest-friendly	12	17	50	21
		24	17	28	31
	ABAA	9	53	21	17
		14	59	16	12
	hate-detest-love-dislike	12	14	53	21
	1-2-3-4	19	28	31	22
	AABA	26	19	40	16
	Ace-2-3-Joker	28	14	29	29
	BAAA	31	24	19	26
	AAAB	16	22	22	40
RepresentRating†		10	32	27	31
	polite-rude-honest-friendly	34	7	26	32
		38	32	27	3
	ABAA	38	5	30	28
		29	19	30	21
	hate-detest-love-dislike	34	15	20	30
	1-2-3-4	33	23	16	28
	AABA	34	23	8	35
	Ace-2-3-Joker	39	15	17	29
	BAAA	11	35	25	29
	AAAB	43	25	24	8
SalienceRating†		94	2	4	0
	polite-rude-honest-friendly	14	57	21	8
		5	6	8	81
	ABAA	2	91	4	4
		13	72	11	4
	hate-detest-love-dislike	16	19	62	3
	1-2-3-4	38	21	25	16
	AABA	3	4	93	1
	Ace-2-3-Joker	28	3	3	66
	BAAA	92	3	5	0
	AAAB	5	6	3	86

†In case a participant rated several items as most representative/most salient, her count would be evenly distributed on all corresponding locations.

Table A.2: Full data from the complementary tasks (relative choice frequencies; for the Rating tasks: relative frequencies of location ranked the highest).

Appendix B Payoff matrices for the different games

		Seeker			
		1st Option	2nd Option	3rd Option	4th Option
Hider	1st Option	(0,1)	(1,0)	(1,0)	(1,0)
	2nd Option	(1,0)	(0,1)	(1,0)	(1,0)
	3rd Option	(1,0)	(1,0)	(0,1)	(1,0)
	4th Option	(1,0)	(1,0)	(1,0)	(0,1)

Table B.1: Hide&Seek game

		Player 2			
		1st Option	2nd Option	3rd Option	4th Option
Player 1	1st Option	(1,1)	(0,0)	(0,0)	(0,0)
	2nd Option	(0,0)	(1,1)	(0,0)	(0,0)
	3rd Option	(0,0)	(0,0)	(1,1)	(0,0)
	4th Option	(0,0)	(0,0)	(0,0)	(1,1)

Table B.2: Coordination game

		Player 2			
		1st Option	2nd Option	3rd Option	4th Option
Player 1	1st Option	(0,0)	(1,1)	(1,1)	(1,1)
	2nd Option	(1,1)	(0,0)	(1,1)	(1,1)
	3rd Option	(1,1)	(1,1)	(0,0)	(1,1)
	4th Option	(1,1)	(1,1)	(1,1)	(0,0)

Table B.3: Discoordination game

		Player 2			
		1st Option	2nd Option	3rd Option	4th Option
Player 1	1st Option	(0,0)	(0,1)	(0,0)	(1,0)
	2nd Option	(1,0)	(0,0)	(0,1)	(0,0)
	3rd Option	(0,0)	(1,0)	(0,0)	(0,1)
	4th Option	(0,1)	(0,0)	(1,0)	(0,0)

Table B.4: To-your-right game