No. 96

Irenaeus Wolff

On the Salience-Based Level-k Model

Research Paper Series
Thurgau Institute of Economics and Department of Economics
at the University of Konstanz

Member of

thurgauwissenschaft

www.thurgau-wissenschaft.ch

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

On the Salience-Based Level-k Model[§]

Irenaeus Wolff

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

Abstract:
In the current literature, there is a lively debate about whether a level-k model can be based on salience to explain behaviour in games with distinctive action labels such as hide-and-seek or discoordination games. This study presents six different experiments designed to measure salience. When based on any of these empirical salience measures, the standard level-k model does not explain hide-and-seek behaviour. Modifying the model such that players follow salience when payoffs are equal, the model fits hide-and-seek data well. However, neither the original nor the modified model account for data from a discoordination game. This holds true even when basing the level-k prediction on participants' own individual salience assessments.

Keywords: ABAA, hide and seek, cognitive hierarchy, strategic reasoning, saliency.
JEL: C72, C91

1 Introduction

Matching-pennies games and their generalisation to multiple actions, dubbed hide-and-seek games, have been well-studied games in game theory from its

[§]I am grateful to my co-authors Lisa Bruttel, Andreas Nicklisch, David Dohmen, Timo Heinrich, Konstantin Hesler, and Simeon Schudy, for their cooperation on projects that produced some of the data I am using here; the latter four also contributed substantially to this paper through uncountable discussions on the level-k explanation for behaviour in hide-and-seek games. I would like to thank Martin Dufwenberg, Urs Fischbacher, Shaun Hargreaves Heap, David Rojo Arjona, Dirk Sliwka, Robert Sugden, Marie-Claire Villeval, Roberto Weber, the lively research group at the Thurgau Institute of Economics (TWI), as well as the participants of the 2013 GfeW Meeting and the 2014 ESA European Meeting for helpful comments and fruitful discussions. I thank Vincent Crawford and Nagore Iriberri for rapidly answering any questions with respect to their paper, as well as for their comments on an earlier version of this paper. Financial support by the University of Konstanz' Young Scholar Fund is gratefully acknowledged.

very beginning (cf., e.g., von Neumann, 1953). Real-life examples abound, from markets in which the brand leader will continue to have the largest revenues as long as it can match the rival products' features, to obvious applications in military, police, and intelligence work. While the standard game-theoretic solution to the generic games is straightforward, experiment participants do not seem to act according to this prediction (see Eliaz and Rubinstein, 2011, for a repeated matching-pennies game, and Rubinstein and Tversky, 1993, as well as Rubinstein et al., 1996, for hide-and-seek games).

In a well-noted paper, Crawford and Iriberri (2007; henceforth CI) show how a salience-based level-k model can account for the observed patterns in hide-and-seek games if we assume a specific salience pattern. Current work by Hargreaves Heap et al. (2014) shows that a level-k model cannot account simultaneously for data from hide-and-seek games, coordination games, and discoordination games all played on the same action-set frame if we assume the same salience pattern for all games. In a comment on this work, Crawford (2014) argues that level-k should not be applied to coordination games because these games fall into the domain of team-reasoning theory (Sugden, 1995). However, if we take out the coordination games in Hargreaves Heap et al.'s study, we can no longer say anything about the descriptive validity of the level-k model. This is the gap the present paper fills.

In this paper, I provide six empirical measures of what is salient and show that none of them is in line with CI's assumption on salience. More importantly, I show that the proposed level-k model no longer predicts behaviour well when based on any of these empirical salience measures. In contrast to CI's proposed model, my estimations suggest that salience influences behaviour directly, *on top of* determining the anchor of players' belief cascades. A simple and plausible model modification taking this influence into account restitutes the remarkable fit of CI's level-k model. However, neither the original model nor its modification can explain data from discoordination games. Most importantly, I show that a level-k model based on empirical salience does not account for the discoordination-game data even when we account for the fact that there is heterogeneity in participants' elicited salience perceptions (so that every participant may have their own level-0).

This paper contributes to a growing literature that finds empirical support for level-k-like thinking in a variety of games.[1] It also contributes to a small but growing literature on how salience shapes behaviour and how this can be incorporated into game-theoretic models.[2] Crawford and Iriberri (2007) do a remarkable job in joining these two branches of the literature. What the present

[1]E.g., Burchardi and Penczynski (2014) or the many papers cited in Crawford et al. (2013).

[2]E.g, Mehta et al. (1994) or Bardsley et al. (2010).

paper shows is that salience influences on behaviour remain a phenomenon that is not as straightforward as it may seem. It is not obvious *a priori* what is salient in the eyes of experimental participants, and it remains to be understood how exactly salience shapes participant behaviour.

The remainder of this paper is organised as follows: Section 2 presents the hide-and-seek game in its archetype version, next to CI's level-k explanation. In Section 3, I present the six salience-elicitation experiments, the results of which are presented in Section 4.1. Section 4.3 presents the equilibrium model with salience-based payoff-perturbations CI use as a benchmark, and Section 4.4 introduces the level-k models used in this paper, including the new modification introduced here. In Section 4.5, I report on the estimates that result for the models when these are based on an experimentally-elicited salience-pattern. In Section 4.6, I explore the level-k model's predictive power in a coordination and a discoordination games. Section 4.7 incorporates a heterogeneous level-0 using participants' own individual salience assessments, and evaluates the resulting models' data fit. Section 5 summarises the data and discusses the findings. An explanation of the model denotations used throughout the paper can be found in Section 4.2.

2 Hide-and-seek and the level-k explanation

In the archetype version of the hide-and-seek game, a "hider" possesses a "treasure" she can hide in one of four boxes, labelled "A", "B", "A", and "A". A "seeker" may open one of these boxes. If he chooses the same box as the hider, the seeker gains the treasure, otherwise the hider keeps it. This multiple-action matching-pennies game obviously has a unique Nash equilibrium in mixed strategies, with both the hider and the seeker choosing each box with 25% probability. The typical distribution observed in experimental implementations of the game, on the other hand, has a strong mode on "central A" for both roles, being even more pronounced for seekers than for hiders (which leads to a substantial seeker-advantage relative to equilibrium).[3]

Let us now turn to how a level-k model may account for the above pattern. Level-k models have a very simple structure. Each k-type, $k > 0$, believes all her opponents are of level-$(k-1)$ and best-responds to this belief.[4] The two

[3]The data from Rubinstein et al.'s experiments are reported in Appendix A.

[4]This type of model was introduced by Stahl and Wilson (1994, 1995) and Nagel (1995), and later adapted by Costa-Gomes et al. (2001). It is closely-related to other *cognitive-hierarchy* models like that proposed by Ho et al. (1998) and refined in Camerer et al. (2004). For a discussion of both approaches, cf. Crawford et al. (2013). Note that Crawford and Iriberri (2007) allow for errors in their model. However, given the estimated error rate for the models under the assumption

crucial elements to close the model are the specifications of level-0—which is assumed to exist in the players' minds only—and of the type distribution. For the latter, CI argue that "[t]he estimated distribution tends to be stable across games and hump-shaped," (p. 1734) while the level-0 specification is the central innovation of their paper. Instead of assuming the traditional uniform mixture over all possible actions, CI 'translate' Rubinstein and co-authors' statements on salience into (latent) numeric variables to use them as level-0 in their model. In CI's words,

> "[t]he 'B' location is distinguished by its label, and so is salient in one of Thomas Schelling's (1960) senses. And the two 'end A' locations, though not distinguished by their labels, may be inherently salient, as RT [Rubinstein and Tversky, 1993] and RTH [Rubinstein, Tversky, and Heller, 1996] argue, citing Nicholas Christenfeld (1995). As RT note, these two saliencies interact to give the remaining location, 'central A,' its own brand of uniqueness as 'the least salient location.' " (p. 1732).

CI translate the last sentence as implying that "central A" really *is* "the least salient location," thus being chosen by a level-0 player least often. I argue that this need not be true. If "central A" has "its own brand of uniqueness", it is not clear *a priori* how it should be ranked in terms of salience. The evidence presented in this paper suggests "central A" is in fact more salient than "final A", whereas it is unclear how it compares to "first A" in terms of salience.

3 Design of the salience-elicitation experiments

The purpose of the salience-elicitation exercise is to provide a clearer understanding of what may constitute an adequate level-0 specification for the model. I argue that there are multiple ways of how salience could determine level-0 that are associated with distinct empirical measures. On the level of beliefs involved, I follow three approaches: a first approach is to define level-0 directly in terms of the available actions' salience (*primary salience* in Bardsley et al., 2010, referring to Lewis, 1969). This corresponds most closely to CI's proposed *model.* A second approach is to ask what people think will be salient for other people (*secondary salience* in Bardsley et al., 2010, also referring to Lewis, 1969). This corresponds more closely to CI's *general reasoning* about level-0, given level-0 is meant to exist only in the players' minds. Finally, we may be tempted to argue that the truly

of uniform errors is zero, I abstract from errors for the time being. None of the findings hinges on this simplification.

relevant aspect would be to ask what people think others will think everybody will hold to be salient, and so on, *ad infinitum*.[5]

It can be argued that only primary salience should be a candidate for level-0, given secondary and higher-order salience involve strategic thinking in the sense of guessing about others' perceptions or even about others' reasoning about these perceptions. Because strategic thinking is what the model should explain, so the argument, we should not include strategic thinking as an input into the theory *via* level-0. In this paper, I choose to be more lenient with the theory by allowing for higher-order salience, to give the model a greater chance of being in accordance with the data. If we cannot accomodate the data, neither when not allowing some strategic deliberation to sneak in through our salience measures nor when doing so, this will be more informative than if we stuck to only the more puristic version of the model. On the other hand, if there were a model able to explain choices based on empirically-measured secondary salience, I would see it as a fruitful first step to a more complete model that gets rid of the problem. I therefore include also secondary and 'infinite-level' salience as candidates for level-0.[6]

Having looked at the above 'levels of salience', I also want to test whether the game description will shape the salience of the available actions. More precisely, players may assess an action's salience differently, depending on whether they look at the actions *per se*, or whether they look at the actions taking into account the game they will be playing.[7] In the latter case, it would be plausible also to assume that players' roles may affect their salience assessment.[8]

In this study, I examine six experimental measures of a salience-based level-0. Note that the point of this exercise is not to compare the different measures. Rather, I want to test whether *any* of these measures would yield a salience pattern that, being plugged into CI's level-k model, would allow that model to account for the data.[9]

The first three salience measures I use are a full variation along the belief dimension, keeping the game description out. The fourth measure uses the *secondary-salience* measure to explore the effect of introducing the game story (and whether an asymmetry follows from that). Measures five and six provide alternative measures of *primary salience* with and without the game story.[10] To be precise, I look

[5]Bardsley et al. (2010) point out there may be higher 'levels of salience' but argue that they are likely to coincide with *secondary salience*. My results would support this conjecture.

[6]I thank Hargreaves Heap et al. for raising this point.

[7]In fact, the study by Hargreaves Heap et al. (2014) suggests this may be the case.

[8]CI partially incorporate this latter aspect by presenting different model specifications, e.g., including a salience-seeking level-0 seeker and a salience-avoiding level-0 hider.

[9]Hence, no care was taken to have similar numbers of observations in the different treatments.

[10]Measures five and six were added because some commenters on an earlier draft raised doubts

at the following experiments:[11]

PICKING TASK. Elicitation of the different boxes' salience by asking people to choose one of four boxes labelled "A", "B", "A", and "A", and click on the chosen box, on a separate page of a post-experimental questionnaire after an unrelated experiment.[12] This is the "picking task" Bardsley et al. (2010) use to elicit *primary salience*. As a crucial complementary measure to assess salience, I record response times for this task.

GUESSING TASK. Elicitation of what people think will be salient for other people. For this purpose, we ask participants to estimate the relative click frequencies from the answers elicited by the PICKING TASK.[13] This is the "guessing task" Bardsley et al. (2010) use to elicit *secondary salience*.

BEAUTY CONTEST. A beauty contest on the consensus on what is salient. The contest anchored in the question "which is the most salient box, which are the second, third, and fourth most salient boxes?" and was conducted as a classroom experiment in the Experimental Methods course.[14]

POST-GAME GUESSING. Elicitation of what people who know the game think will be salient for others. For this purpose, we asked participants to estimate the relative click frequencies of the PICKING TASK responses. This was done *after* they had played the hide-and-seek game but before they got any feedback. This measure serves as a benchmark for how the game—and possibly, the role—changes salience-perceptions.[15]

RATING TASK. Participants were asked to rate the salience of each of the four boxes on an 11-point Likert scale ranging from "extremely inconspicuous" to "extremely conspicuous".

POST-STORY RATING. Participants were explained the hide-and-seek game in a role-neutral format. Then, they completed the RATING TASK. They did not play the game itself.

---

about the construct validity of the PICKING TASK used in earlier studies (such as Mehta et al., 1994, or Bardsley et al., 2010), even if augmented by response times like in our first measure.

[11] A translated version of the instructions to each task is provided in Appendix B.

[12] The post-experimental questionnaire mainly contains questions from the 16PF personality inventory. Participants have not participated in any hide-and-seek experiment before.

[13] The task was incentivised in the following way: if no frequency differed from the true value by more than 5% (10%/20%), participants could earn an additional 50 (25/10) Euro cents, otherwise, they did not earn anything. The task was the first task participants faced in the experiment, they knew there would be further tasks, but they did not know what those tasks would be.

[14] Amongst those stating the modal ordering, a prize of 12 Euros (about USD 15.60 at the time) was raffled off.

[15] Incentives as in the GUESSING TASK.

None of the participants participated in more than one of the six experiments. The first four Experiments were run at the University of Konstanz, the Picking Task, the Guessing Task, and the Post-Game Guessing at its *Lakelab*. The Rating Task and the Post-Story Rating were added as a questionnaire to a completely unrelated study run at the University of Hamburg. There have not been other studies using an ABAA-like setup at the University of Hamburg.

4 Results

This section is organised as follows: first, I report on the outcomes of the salience-elicitation exercises. Section 4.2 introduces the denominations of all models that appear in the paper. Section 4.3 briefly reviews what the elicited salience-measures mean for Crawford and Iriberri's (2007; CI) benchmark model of an equilibrium with payoff perturbations. Section 4.4 presents the level-k models used, including the modification introduced in this paper. Section 4.5 looks at what the salience-elicitation exercises would mean for the different models, by replicating CI's model-fitting exercise using the elicited salience patterns as level-0. Section 4.6 explores whether any of the level-k models can predict data from a coordination and a discoordination game. Finally, Section 4.7 analyses whether a level-k model based on participants' *individual* salience assessments can account for the discoordination-game data from Section 4.6. For ease of notation, in the remainder of this article I will describe the locations "A", "B", "A", and "A" by $A_{(1)}$, $B_{(2)}$, $A_{(3)}$, and $A_{(4)}$, respectively.

4.1 Salience in the ABAA hide-and-seek game

The results of the six salience-elicitation experiments are reported in Table 1, together with the respective numbers of independent observations.

Observation 1. $B_{(2)}$ is the most salient alternative, and $A_{(3)}$ is not the (single) least salient alternative.

The first part can be seen easily by looking at the second and third data columns in Table 1. Treating the different salience measures as independent realisations of an underlying 'true' salience pattern, we can construct the crudest-possible statistical measure as follows: Assuming that the next-salient candidate has an equal chance of coming out as the most salient alternative on each of the seven measures (frequency and response time in the Picking Task, estimate by Guessing-Task participants and by hiders and by seekers in Post-Game Guessing, any ranking measure in the Beauty Contest), and the ratings in the Rating Task and in Post-Story Rating, we can compute the p-value of the according

	$A_{(1)}$	$B_{(2)}$	$A_{(3)}$	$A_{(4)}$
PICKING TASK (405 participants)				
relative click frequencies (in %)	21	38	35	6
response times (in sec)				
mean	8.8	7.7	8.5	11.9
median	8.0	7.1	7.5	9.4
GUESSING TASK (72 participants)				
average estimated relative click frequency	21	41	22	15
BEAUTY CONTEST (30 participants)				
rank in beauty contest				
winning order (chosen by 14 participants)	2	1	3	4
mean ranks	2.3	1.5	2.5	3.6
POST-GAME GUESSING (156 participants)				
average estimated relative click frequency	19	39	24	18
...by hiders (78 obs.)	19	38	24	19
...by seekers (78 obs.)	19	40	25	17
RATING TASK (90 participants)				
average conspicuousness reported (0 to 10)	5.7	7.5	5.6	5.3
POST-STORY RATING (90 participants)				
average conspicuousness reported (0 to 10)	3.8	7.4	4.3	4.0

Table 1: Salience assessments of the four boxes denoted by "A", "B", "A", and "A".

binomial test to be $p = 1/128$. The same line of argument yields that $A_{(3)}$ is more salient than $A_{(4)}$, with the same level of significance.

Observation 2. From the six different salience measures, I extract three possible salience-patterns: $B_{(2)}[A_{(3)}A_{(1)}]A_{(4)}$, $B_{(2)}A_{(3)}[A_{(1)}A_{(4)}]$, and $B_{(2)}[A_{(3)}A_{(1)}A_{(4)}]$ (locations ordered by salience, square brackets bundle equally-salient locations).

The first pattern, $B_{(2)}[A_{(3)}A_{(1)}]A_{(4)}$, can be observed in the GUESSING TASK, arguably in the mean ranks of the BEAUTY CONTEST, and possibly in the RATING TASK. The second pattern, $B_{(2)}A_{(3)}[A_{(1)}A_{(4)}]$, can be seen in POST-GAME GUESSING and possibly in POST-STORY RATING, while both the RATING TASK and the POST-STORY RATING data can be interpreted as yielding the pattern $B_{(2)}[A_{(3)}A_{(1)}A_{(4)}]$. In addition, one might argue that the PICKING TASK yields $[B_{(2)}A_{(3)}]A_{(1)}A_{(4)}$, but the response times clearly indicate that $B_{(2)}$ and $A_{(3)}$ are salient to different degrees.[16] Note also that for the predictions of CI's level-k model, the potential patterns $B_{(2)}A_{(3)}A_{(1)}A_{(4)}$ (median response times in the PICKING TASK) or

[16]None of the conclusions in this paper would change if we included $[B_{(2)}A_{(3)}]A_{(1)}A_{(4)}$ in the list of salience patterns.

$B_{(2)}A_{(1)}A_{(3)}A_{(4)}$ (winning order of the Beauty Contest) are equivalent to the pattern $B_{(2)}[A_{(3)}A_{(1)}]A_{(4)}$.

Observation 3. The description of the game affects salience but does not lead to any reversal in the rank order of the locations' salience ranks. Moreover, participants' roles in the hide-and-seek game do not seem to influence their estimates of other people's salience perceptions.

For the first part of Observation 3, note that including the description of the game before the Rating Task has average reported conspicuousness of all A-alternatives drop sharply ($p < 0.001$ for $A_{(1)}$, $p = 0.993$ for $B_{(2)}$, and $p = 0.004$ for $A_{(3)}$ and $A_{(4)}$). Further, compare the average estimated relative click frequencies of the Guessing Task and Post-Game Guessing. While the largest quantitative difference between the corresponding average estimates in the Guessing Task and Post-Game Guessing is a mere 3%, there seems to be a clear difference in the qualitative pattern: in the Guessing Task, there is a substantial difference between the average estimated relative click frequency of $A_{(1)}$ and $A_{(4)}$ (and none between $A_{(1)}$ and $A_{(3)}$), while in Post-Game Guessing, the average estimates for $A_{(1)}$ and $A_{(4)}$ are virtually identical (and there is a clear difference with respect to $A_{(3)}$).[17] At the same time, no two locations that would be differently salient in one direction by a Post-Game/Post-Story measure are differently salient in the other direction by the corresponding measure in which participants do not know the game. For the second part of Observation 3 note that within the Post-Game Guessing measure, the qualitative pattern clearly is the same for hiders and seekers, and the quantitative difference between the average estimates is 2% at most.

4.2 Preliminaries: model denominations

Throughout this paper, I will work with a variety of models to account for behaviour. Table 2 is meant to systemize them sufficiently so that it is easier to refer to the different models in the text. There are two main aspects on which the models differ: the salience-pattern on which the model is based, and on the hypothesized strategic thinking given this salience-pattern. In terms of the latter, I will refer to three kinds of models: naïve responses driven by salience that do not require any strategic thinking at all; equilibrium (eqm) models, potentially

[17] Wilcoxon matched-pairs signed-ranks tests support these observations: in the Guessing Task, they yield $p \leq 0.001$ for the comparisons of a participant's $A_{(4)}$-estimate with both her $A_{(1)}$-estimate and her $A_{(3)}$-estimate, while for the comparison of her $A_{(1)}$-estimate with her $A_{(3)}$-estimate, the test yields $p = 0.883$. In Post-Game Guessing, the same test yields $p = 0.133$ for the comparison between $A_{(1)}$ and $A_{(4)}$, and $p < 0.001$ for the comparisons between $A_{(3)}$ and both $A_{(1)}$ and $A_{(4)}$.

Model of strategic thinking	
NAÏVE	Players choose according to salience, no strategic thinking involved.
EQM_0	Assumes rationality and common knowledge thereof; no payoff perturbations.
EQM_+	Assumes rationality and common knowledge thereof; there are payoff perturbations that follow salience.
Lk	Each level-k player best-responds to a level-($k-1$) player; when the best-response is not unique, players randomise uniformly over all best-responses.
Lk_{MOD}	Each level-k player best-responds to a level-($k-1$) player; when the best-response is not unique, players randomise according to the best-responses' salience; when a level-k' player randomises, a level-$k'+1$ player best-responds to a level-k' player's true mix.
Salience-pattern origin	
HYP	The salience pattern is inferred by model-fitting.
NEUT	The salience pattern used is measured 'neutrally', that is, without participants knowing about any of the games.
RATE	The salience pattern used stems from the RATING TASK.
POSTRATE	The salience pattern used stems from POST-STORY RATING.
POSTX	The salience pattern used stems from POST-GAME GUESSING after participants have played game X. X can be H&S for the hide-and-seek, COORD for the coordination, and DISCOORD for the discoordination game.
Salience-pattern used	
W[XY]Z	This postfix repeats the salience ranking used in the model. Locations W to Z are ordered by decreasing salience, square brackets indicate indifference. In the example, location W is the most, and Z the least salient location, while X and Y are equally salient locations.
INDL0	This postfix means the model predictions use participants' individual salience measurements as their respective level-0.
AVOID	This additional postfix indicates that players are assumed to use a salience-avoiding level-0. It is used only in two specifications provided in Appendix D for completeness (see ftn. 23).
ASYM	This additional postfix indicates that hiders (seekers) are assumed to use a salience-avoiding(-loving) level-0. It is used only in two specifications provided in Appendix D for completeness (see ftn. 23).

Table 2: Systemization of the models used in this paper.

including salience-based payoff perturbations; and level-k (Lk)-models. In terms of the salience-pattern used, I will refer to the hypothesised pattern in Crawford and Iriberri (2007; HYP-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$), the data from the PICKING-TASK, GUESSING TASK, and BEAUTY CONTEST (NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$), to the pattern from the (POST-STORY) RATING TASK (RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$), and from different POST-GAME GUESSING experiments (POSTX, where X is a wildcard referring to the respective game). To give an example, Lk-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ denotes a level-k model that is based on participants' salience assessment as elicited in a guessing task conducted after a hide-and-seek game.

		Seeker			
		$\mathbf{A}_{(1)}$	$\mathbf{B}_{(2)}$	$\mathbf{A}_{(3)}$	$\mathbf{A}_{(4)}$
Hider	$\mathbf{A}_{(1)}$	$1-e$ $0+e$	$0+f$ $1+e$	0 $1+e$	$0-e$ $1+e$
	$\mathbf{B}_{(2)}$	$0-e$ $1-f$	$1+f$ $0-f$	0 $1-f$	$0-e$ $1-f$
	$\mathbf{A}_{(3)}$	$0-e$ 1	$0+f$ 1	1 0	$0-e$ 1
	$\mathbf{A}_{(4)}$	$0-e$ $1+e$	$0+f$ $1+e$	0 $1+e$	$1-e$ $0+e$

Table 3: The hide-and-seek game with payoff perturbations when $A_{(1)}$ and $A_{(4)}$ are equally salient (adapted from Crawford and Iriberri, 2007, Figure 2).

4.3 Crawford and Iriberri's (benchmark) equilibrium model with payoff perturbations

In the following section, I briefly present the model of an equilibrium with hard-wired payoff perturbations CI use as a benchmark. CI start with the normal form game and posit that players will have a preference for some locations which depends on those locations' salience. Hiders are assumed to dislike choosing salient locations, while seekers are assumed to favour them. Here, I use the salience measure POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$: assume hiders will obtain an extra benefit (seekers incur a cost) of e when they choose one of the end locations and a cost (a benefit) of f when they choose $B_{(2)}$. If $A_{(1)}$ and $A_{(4)}$ are jointly least salient (as in POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$), we should expect $e > 0$ and $f > 0$.[18] Table 3 shows the resulting normal form. Using RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$ results in the same normal form with the additional restriction that $e \equiv 0$.

Alternatively, we can base the game with payoff perturbations also on NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$. This specification fits the data worse and is therefore relegated to Table C.3 in Appendix C.

4.4 Level-k models used

Just like CI base their model on the salience pattern HYP-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$, we can now base a level-k model on the elicited salience measures NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$,

[18] Given CI posit that $A_{(1)}$ and $A_{(4)}$ are jointly *most* salient, they write down the model using $e' \equiv -e$ and expect the maximum-likelihood estimation to yield $e' > 0$. However, given CI do not put restrictions on the signs of e and f (EQM$_{+}$-HYP-$\left[A_{(1)}A_{(4)}\right]$), both games are equivalent.

POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$, and RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$.[19] Columns three and four of Table 4 present players' predicted choices depending on their k-level exemplarily for NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$.[20]

In addition, I suggest a slight modification of the model, denoted by Lk_{MOD}. Mehta et al. (1994) argue that players who do not have any incentives to favour one action over another will choose by the action labels' salience. If a player is indifferent between multiple actions, she does not have an incentive to favour any of these actions. Hence, it would be very natural to assume that a player who should be indifferent between multiple actions randomises over these actions according to their relative salience rather than randomising uniformly as in standard game theory. This may happen for a variety of different reasons: different people might be inherently attracted to different locations when no compelling economic force acts on them; they might decide to choose 'just anything' from among the options they are indifferent about, in a similar fashion as participants in our PICKING TASK will have chosen one of the boxes when there was no reason to favour any box over the other; or they might try to randomise uniformly, but the attraction exerted by salience might unconsciously interfere with their randomisation attempts. To incorporate this idea, I have to make an additional assumption: a level-$(i+1)$ player is aware of the randomising level-i player's inability not to be attracted by salience, and best-responds to the resulting probability distribution.[21] Columns five and six of Table 4 present the resulting choice predictions exemplarily for POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$.

4.5 Model fit under the elicited salience patterns

Using the same data as CI, I perform a complete grid search over all possible type-distributions (at the percent level), to find the level-k distribution under which the data has the highest log-likelihood, using equation (2) in CI.[22] Table 5 presents the results. The focal models in Table 5 are CI's preferred level-k model using

[19] As an alternative specification, CI estimate their models also for the salience pattern HYP-$B_{(2)}\left[A_{(1)}A_{(4)}\right]A_{(3)}$. Given that HYP-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$ yields the better fit, however, they accept the latter as the pattern to base their preferred model on.

[20] The hide-and-seek data to be fitted is reproduced in Appendix A.

[21] The alternative assumption, in which higher-level players are unaware of salience-influences on randomisation by lower-level players, is explored in the working-paper version Wolff (2014) and yields a worse fit to the data.

[22] Note that I present this analysis primarily for comparability. By the logic of this paper, I should restrict myself to a small subset of the data: CI use data from 6 different treatments conducted by Rubinstein and co-authors; to use all 6, CI have to make assumptions of how to convert the data from some treatments to make them comparable to the data from others. E.g., data from a treatment using "A", "A", "B", and "A" is adapted by simply switching the two locations in the middle. Whether this is appropriate based on the underlying salience structure is, again, an empirical question. I refer to the analysis of Heinrich and Wolff's (2012) data in the right-hand

k-level (frequency)	box	Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$ Hider	Seeker	Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ Hider	Seeker
L0 (π_0)	A(1)	0.21	0.21	0.19	0.19
	B(2)	0.38	0.38	0.39	0.39
	A(3)	0.35	0.35	0.24	0.24
	A(4)	0.06	0.06	0.18	0.18
L1 (π_1)	A(1)	0	0	0.51	0.00
	B(2)	0	1	0.00	1.00
	A(3)	0	0	0.00	0.00
	A(4)	1	0	0.49	0.00
L2 (π_2)	A(1)	1/3	0	0.31	0.51
	B(2)	0	0	0.00	0.00
	A(3)	1/3	0	0.39	0.00
	A(4)	1/3	1	0.30	0.49
L3 (π_3)	A(1)	1/3	1/3	0.00	0.00
	B(2)	1/3	0	0.62	0.00
	A(3)	1/3	1/3	0.38	1.00
	A(4)	0	1/3	0.00	0.00
L4 (π_4)	A(1)	0	1/3	0.25	0.00
	B(2)	1	1/3	0.51	1.00
	A(3)	0	1/3	0.00	0.00
	A(4)	0	0	0.24	0.00
Total	A(1)	$0.21\pi_0 + \frac{\pi_2+\pi_3}{3}$	$0.21\pi_0 + \frac{\pi_3+\pi_4}{3}$	$0.19\pi_0 + 0.51\pi_1 + 0.31\pi_2 + 0.25\pi_4$	$0.19\pi_0 + 0.51\pi_2$
	B(2)	$0.38\pi_0 + \frac{\pi_3}{3} + \pi_4$	$0.38\pi_0 + \pi_1 + \frac{\pi_4}{3}$	$0.39\pi_0 + 0.62\pi_3 + 0.51\pi_4$	$0.39\pi_0 + \pi_1 + \pi_4$
	A(3)	$0.35\pi_0 + \frac{\pi_2+\pi_3}{3}$	$0.35\pi_0 + \frac{\pi_3+\pi_4}{3}$	$0.24\pi_0 + 0.39\pi_2 + 0.38\pi_3$	$0.24\pi_0 + \pi_3$
	A(4)	$0.06\pi_0 + \pi_1 + \frac{\pi_2}{3}$	$0.06\pi_0 + \pi_2 + \frac{\pi_3}{3}$	$0.18\pi_0 + 0.49\pi_1 + 0.3\pi_2 + 0.24\pi_4$	$0.18\pi_0 + 0.49\pi_2$

Table 4: Players' hide-and-seek choice probabilities under Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$ and Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$. The fraction of players of level i is denoted by π_i.

the empirically-elicited salience patterns as level-0, Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$, L$k$-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$, and L$k$-POSTH&S-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$, as well as the modified model Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$.[23] For comparison, I include also estimates of the following seven benchmark models: choice according to the empirically-elicited salience patterns (NAÏVE-PICK-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$, NAÏVE-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$, NAÏVE-RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$); the standard mixed-strategy Nash-equilibrium (EQM$_0$); CI's equilibrium with "unrestricted" payoff perturbations based on $A_{(1)}$ and $A_{(4)}$ being equally salient (EQM$_+$-HYP-$\left[A_{(1)}A_{(4)}\right]$), as well as with 'partially restricted' perturbations (so as to match the

part of Table 5 for an analysis that does not rely on comparable assumptions.

[23] CI's alternative level-k specifications with an asymmetric level-0 (favouring salience for seekers and avoiding it for hiders, Lk-X-ASYM) and with a salience-avoiding level-0 (Lk-X-AVOID) are included in Table D.4 in Appendix D.

elicited salience pattern; EQM$_+$-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$);[24] CI's preferred level-k model under their salience assumption (Lk-HYP-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$).

Readers may object that the salience-elicitation experiments where all conducted in Germany, and therefore, in a different cultural environment than the actual games. Furthermore, if people from different cultures have different perceptions in terms of salience or if their salience-based strategic reasoning is shaped culturally, we cannot conclude much from elicitating salience in one part of the world to explain behaviour in another. To respond to this valid objection, I also include the model estimates for German hide-and-seek data, taken from a study by Heinrich and Wolff (2012).[25]

Observation 4. 'Random clicking' as shaped by salience patterns does not explain the data well.

Observation 4 rests on the fact that both the log-likelihoods and the mean squared errors of all three NAÏVE-specifications indicate a fit that is even worse than the equilibrium prediction without payoff perturbations. This is important because naïve, unstrategic responses are one of two explanations for hide-and-seek data in the literature.

Main Result 1. Using the same data as Crawford and Iriberri (2007), the best measured-salience-based estimates for their preferred level-k model fit the data clearly worse than the estimates they derive for an equilibrium model with 'unrestricted' payoff perturbations.

Main Result 1 can be verified by a look at the Table-5 columns reporting the log-likelihoods, comparing specification Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$ to specification EQM$_+$-HYP-$\left[A_{(1)}A_{(4)}\right]$.[26] Note that it does not depend on the level-k dis-

[24]The estimates of EQM$_+$-HYP-$\left[A_{(4)}\right]$ (no restriction on the sign of e) and EQM$_+$-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$) are also included in Table D.4 in Appendix D. The estimate for EQM$_+$-RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$ coincides with the one for NAÏVE-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ and hence is omitted. For all EQM$_+$-estimations, I use a two-step procedure: I first do a complete grid search over all four parameters for $-1 \leq e_H, f_H, e_S, f_S \leq 1$ at the five-percent level, and then another one at the percent level for the parameter space $[e_H - 0.1, e_H + 0.1] \times [f_H - 0.1, f_H + 0.1] \times [e_S - 0.1, e_S + 0.1] \times [f_S - 0.1, f_S + 0.1]$.

[25]For comparability, I include only the data obtained under the original instructions. As pointed out in footnote 22, this data has the additional advantage that it was obtained exclusively under the ABAA-protocol, so that no further assumptions are needed of how to translate salience patterns from other setups, such as the AABA-protocol.

[26]The result holds also for all other level-k variants presented in CI: both Lk-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$-ASYM and -AVOID exhibit log-likelihoods of -1603 (RTH's data) and -465 (HW's data), cf. Table D.4 in Appendix D. Further, it holds for an alternative Lk-NEUT-$\left[B_{(2)}A_{(3)}\right]A_{(1)}A_{(4)}$ specification that Hargreaves Heap et al. have suggested would fit the PICKING-TASK data better (logL of -1643, RTH's data, and of -471, HW's data). I am not pre-

Specification	RTH's data logL	RTH's data MSE	HW's data logL	HW's data MSE
Choices follow salience				
NAÏVE-PICK-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$	-1724	0.01271	-521	0.01654
NAÏVE-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	-1687	0.01647	-487	0.01662
NAÏVE-RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	-1663	0.01226	-486	0.01525
Equilibrium models				
EQM0	-1641†	0.00967†	-484	0.01436
EQM+-HYP-$\left[A_{(1)}A_{(4)}\right]$	-1562†	0.00006†	-456	0.00109
($e_H=-0.29, f_H=0.25, e_S=-0.15, f_S=0.15)^{\dagger}$				
EQM+-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	-1636	0.00909	-483	0.01467
($e_H=0.00, f_H=0.06, e_S=0.00, f_S=0.05$)				
CI's preferred model				
Lk-HYP-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$	-1564†	0.00027†	-456	0.00109
Lk-NEUT-$\mathbf{B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}}$	**-1616**	**0.00683**	**-476**	**0.01192**
Lk-POSTH&S-$\mathbf{B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]}$	**-1635**	**0.00854**	**-485**	**0.01514**
Lk-RATE-$\mathbf{B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}}$	**-1629**	**0.00830**	**-480**	**0.01259**
Modified level-k model				
Lk_{MOD}-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$	-1597	0.00310	-457	0.00152
best 'hump-shaped' type distribution			-458	0.00166
Lk$_{\text{MOD}}$-POSTH&S-$\mathbf{B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]}$	**-1570**	**0.00097**	**-458**	**0.00143**
Lk_{MOD}-POSTRATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	-1621	0.00734	-477	0.01150

† indicates the estimate is taken from CI's paper. ‡ The better-performing specification from RATING TASK and POST-STORY RATING.

Table 5: Log-likelihoods and mean squared errors of the maximum-likelihood estimates of the indicated models. The first two data columns use the data from Rubinstein, Tversky, and Heller's collected studies ("RTH"), reproduced in Table 3 of Crawford and Iriberri (2007; "CI"). Columns three and four replicate the findings using Heinrich and Wolff's (2012; "HW") data. The data from both studies is provided in Appendix A.

tribution we use—that is, it holds even for the estimates yielding the highest likelihoods. These distributions are depicted in Table 6.

Observation 5. A maximum-likelihood estimate of the best-performing elicited-salience-based variant of Crawford and Iriberri's (2007) preferred model yields a level-k distribution that is U- rather than hump-shaped. At the same time, it indicates substantial levels of level-0.

The first part of Observation 5 follows from the row corresponding to specification Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$ in Table 6. As was already stated, "[t]he estimated distribution tends to be stable across games and *hump-shaped*," (CI, p. 1734, emphasis added) which renders the estimate implausible. For the second

senting this alternative in the main text because the measured reaction times in my view suggest $B_{(2)}$ and $A_{(3)}$ are salient to different degrees—which is supported by the RATING-TASK data.

Specification	RTH's data					HW's data				
	$L0$	$L1$	$L2$	$L3$	$L4$	$L0$	$L1$	$L2$	$L3$	$L4$
Lk-HYP-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$	0.00†	0.19†	0.32†	0.24†	0.25†	0.00‡	0.12	0.37	0.29	0.22
Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$	0.38	0.14	0.00	0.48	0.00	0.59	0.14	0.08	0.19	0.00
Lk-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	0.09	0.28	0.01	0.62	0.00	0.70	0.07	0.00	0.23	0.00
Lk-RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$	0.56	0.00	0.44	0.00	0.00	0.61	0.00	0.33	0.06	0.00
Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	0.00	0.20	0.36	0.44	0.00	0.00	0.22	0.28	0.50	0.00

† indicates the estimate is taken from CI's paper. ‡ As CI do not rely on any specific $L0$-pattern, it is not clear to me how to endogenise π_0 without including yet another two parameters (p and q in their paper). As I am reluctant to do so, I maintain $\pi_0 \equiv 0$ for their preferred model.

Table 6: Level-k distributions of the maximum-likelihood estimates in Table 5. The first four data columns use the data from Rubinstein, Tversky, and Heller's collected studies ("RTH"), reproduced in Table 3 of Crawford and Iriberri (2007; "CI"). Columns five to eight replicate the findings using Heinrich and Wolff's (2012; "HW") data.

part of Observation 5, note that the model estimates the level-0 fraction to be 38% or even higher.

Main Result 2. The modified level-k model is able to fit the data substantially better compared to Crawford and Iriberri's (2007) level-k variant when the latter also is based on empirically-elicited salience patterns. The best fit to the data—being almost as good as the fit of the equilibrium with unrestricted perturbations—is achieved by the model in which level-0 is given by the data from the POST-GAME GUESSING task.

The first claim rests on a comparison of the log-likelihoods of Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ in Table 5 to those of Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$, L$k$-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$, and L$k$-RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$.[27] For the second, note that the log-likelihood of Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ in Table 5 is very close to that of EQM$_+$-HYP-$\left[A_{(1)}A_{(4)}\right]$ in the same table. Furthermore, the estimated level-distributions in Table 6 indicate that the best-fitting modified model does exhibit a—albeit skewed—hump-shaped levels distribution. Observation 6 points out a likely reason for the good performance of the Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ model:

Observation 6. The measured salience pattern influences behaviour directly, on top of being the anchor for players' belief cascades.

Observation 6 emphasises the fact that the salience patterns that I elicited and plugged into the level-k models capture an important part of behaviour.

[27] The claim also holds true with respect to CI's other level-k variants, cf. Table D.4 in Appendix D.

In CI's preferred model, salience enters only through *virtual* level-0 players; when basing the model on our salience measures, however, virtually all estimates (and particularly, the best-performing one) indicate a large fraction of—salience-guided—level-0 play. On the other hand, when salience enters the randomisation of otherwise indifferent players, as in the modified model Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$, allowing for level-0 in the estimation has no additional explanatory power.

Finally, note that CI's estimation of the equilibrium with perturbations (EQM$_{+}$-HYP-$\left[A_{(1)}A_{(4)}\right]$) also suggests a salience pattern that exhibits $A_{(3)}$ as the least salient alternative ($e_H, e_S < 0, f_H, f_E > 0$). This would imply that either all of our empirical estimates of salience are wrong or CI's benchmark equilibrium model with payoff perturbations rests on implausible assumptions, too.

Observation 7. Estimates for the equilibrium models with payoff perturbations under the constraint that the payoff perturbations follow any of the elicited salience-pattern candidates have a similarly bad model fit as the re-estimated CI models.

This observation follows from looking at the row in Table 5 pertaining to specification EQM$_{+}$-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$, and comparing the log-likelihood to those of the specifications Lk-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$ and Lk-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$.[28] Note that Observations 4-7 also hold for the data from Heinrich and Wolff (2012), as can be verified by looking at the respective columns in the right-hand part of Tables 5 and 6. The similarity of the estimated parameters and of the models' relative likelihoods suggest that the hide-and-seek game is played in a similar fashion in Stanford, Tel Aviv, and Konstanz.

4.6 Elicited-salience-based level-k in (dis-)coordination games

Hargreaves Heap et al. (2014) show that it is not possible to find a common level-0 that would allow to predict simultaneously behaviour in hide-and-seek, coordination, and discoordination games played on the same action sets with the same labels. In a comment on their work, Crawford (2014; p.5) argues that level-k is the wrong model to account for behaviour in coordination games, given in coordination games "both intuition and existing evidence point toward 'team reasoning'" instead. However, if we take out coordination games from Hargreaves Heap et al. (2014), we can no longer draw the conclusions they draw. Furthermore, even if we did not exclude the coordination-game data, allowing for a different level-0 in each game could be acceptable in principle, given the game description seems to

[28]The same holds true for EQM$_{+}$-RATE-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$ (estimate as for EQM$_{+}$-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$) and for EQM$_{+}$-NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$, cf. Table D.4 in Appendix D.

	$A_{(1)}$	$B_{(2)}$	$A_{(3)}$	$A_{(4)}$
POSTCOORD TASK (72 participants)				
average estimated relative click frequency	19	50	18	14
POSTDISCOORD TASK (72 participants)				
average estimated relative click frequency	20	37	24	19

Table 7: Salience assessments of the four boxes denoted by "A", "B", "A", and "A". The POSTCOORD TASK is the GUESSING TASK after participants played the coordination game, the POSTDISCOORD TASK the same task after participants played the discoordination game, again before any feedback was given.

have a (limited) influence participants' perceptions of salience (Observation 3). Yet, this needs to be done in an objective manner. For this purpose, I repeat the salience-elicitation measure from the best-performing model Lk_{MOD}-POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$. That is, I conduct another GUESSING TASK after participants have played the coordination (POSTCOORD TASK; included for completeness) or discoordination game (POSTDISCOORD TASK), respectively, without feedback.[29] Table 7 shows the results.

Observation 8. The POSTCOORD and POSTDISCOORD TASKS confirm Observation 3: the game description affects participants' estimate of what people in a non-strategic situation will regard as salient, but it does not change the pattern dramatically. In particular, no additional salience pattern is observed.

Observation 8 becomes obvious from Table 7 by focusing on the estimates for locations $B_{(2)}$ and $A_{(4)}$. While in the POSTDISCOORD TASK, the latter is virtually identical to players' estimate on $A_{(1)}$, there is a clear difference in the POSTCOORD TASK.[30] At the same time, the average estimate on $B_{(2)}$ is clearly higher in the POSTCOORD TASK compared to the POSTDISCOORD TASK.[31] Note further that the salience-ranking of the POSTCOORD TASK is the same as NEUT-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$, and that the ranking of the POSTDISCOORD TASK is the same as POSTH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$.

Given the above eliticed salience patterns, we can now predict behaviour in the coordination and discoordination games. To do so, we use the level-k distri-

[29] All procedures as in POST-GAME GUESSING. No participant had participated in any other of the experiments described in this paper.

[30] Two-sided Wilcoxon matched-pairs signed-ranks tests yield $p = 0.174$ and $p = 0.001$, respectively.

[31] A two-sided Wilcoxon Mann-Whitney test yields $p < 0.001$.

butions estimated in Section 4.5.[32] Table 8 reports the predictions of Crawford and Iriberri's preferred model Lk-HYP-$[A_{(1)}A_{(4)}]\,B_{(2)}A_{(3)}$ and the modified model based on the POSTCOORD-TASK and POSTDISCOORD-TASK data, respectively, next to the choices in the coordination and discoordination games proper.

	A$_{(1)}$	B$_{(2)}$	A$_{(3)}$	A$_{(4)}$	MSE
COORDINATION GAME (72 participants)					
Choices (in %)	18	69	11	1	
Lk-HYP-$[A_{(1)}A_{(4)}]\,B_{(2)}A_{(3)}$					
Prediction without errors	50	0	0	50	0.20768
Lk_{MOD}-POSTCOORD-$B_{(2)}\,[A_{(1)}A_{(3)}]\,A_{(4)}$					
Prediction without errors	0	100	0	0	0.03518
Lk_{MOD}-POSTCOORD-INDL0					
Prediction of ML estimate† ($\pi_0 = 0.44$)	14	66	13	7	0.00169
DISCOORDINATION GAME (72 participants)					
Choices (in %)	15	26	40	18	
Lk-HYP-$[A_{(1)}A_{(4)}]\,B_{(2)}A_{(3)}$					
Prediction without errors	28	22	22	28	0.01523
Lk_{MOD}-POSTDISCOORD-$B_{(2)}A_{(3)}\,[A_{(1)}A_{(4)}]$					
Prediction without errors	14	44	28	14	0.01213
Lk_{MOD}-POSTDISCOORD-INDL0					
Prediction of ML estimate ($\pi_0 = 0.91$)	20	37	24	19	0.01003

†Participant 49 excluded in the estimation (choice of a 0-probability event).

Table 8: Choices in and predictions for the coordination and discoordination games. The predictions rest on the estimated fractions of level-k types reported in Table 5. For the modified model allowing for a heterogeneous level-0 (-INDL0), I report the maximum-likelihood estimate, where π_0 is the estimated fraction of level-0 players.

Main Result 3. Neither Crawford and Iriberri's (2007) preferred model nor the modified variant proposed in this paper predicts well the coordination-game and discoordination-game data. The former predicts poorly even qualitatively in both the coordination and the discoordination games, while the latter predicts the qualitative pattern in the coordination game but not in the discoordination game.

To see this, note that the modal choice in the coordination game is $B_{(2)}$, while CI's preferred model Lk-HYP-$[A_{(1)}A_{(4)}]\,B_{(2)}A_{(3)}$ would predict that only players

[32]For the modified level-k model, I use the estimate based on Heinrich and Wolff's (2012) data because that data comes exclusively from the A-B-A-A setup.

making an error would choose this location. Similarly, Lk-HYP-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$ with a level-k distribution as estimated in the hide-and-seek game predicts that in the discoordination game, most participants choose $A_{(1)}$ and $A_{(4)}$, when in fact, the modal choice is $A_{(3)}$ (and $A_{(1)}$ and $A_{(4)}$ are chosen least often). The modified model predicts too concentrated a choice distribution in the coordination game (all choices vs. 69% on $B_{(2)}$), and the wrong modal choice in the discoordination game (44% on $B_{(2)}$ vs. 40% on $A_{(3)}$, while correctly predicting low choice-frequencies for $A_{(1)}$ and $A_{(4)}$).

Main Result 3 establishes that none of the level-k models can account for the data of all three experiments even when we acknowledge that the game descriptions change participants' salience perceptions. What is noteworthy is that the coordination-game data could easily be reconciled with a level-k model with errors or positive fractions of level-0 play—contrary to the conjecture of Crawford (2014)—but the discoordination-game data cannot. Yet, it is the discoordination game where level-k should apply. However, also note that up to now, we stuck to the assumption that all players have the same idea of what is salient. This is empirically wrong. Looking at the POSTDISCOORD TASK as an example, only 60% of the participants estimate that $B_{(2)}$ is clicked on most often in the PICKING TASK, followed by 19% for $A_{(3)}$ and 10% for each $A_{(1)}$ and $A_{(4)}$. Taking this seriously calls for a model that allows every player to have their own level-0. I look at this possibility exemplarily for the coordination and discoordination games in the following section.

4.7 Level-k based on participants' individual salience assessments

In the final lines of each part in Table 8, I report maximum-likelihood estimates for the modified level-k model when participants use their respective own individual POSTCOORD/POSTDISCOORD TASK responses as level-0.[33] To give an example, assume that a participant in the POSTDISCOORD TASK estimates responses in the PICKING TASK to follow the distribution 10%, 45%, 15%, and 30% for $A_{(1)}$, $B_{(2)}$, $A_{(3)}$, and $A_{(4)}$, respectively. In that case, the prediction for the participant's behaviour in the discoordination game would be that she chooses $A_{(1)}$ for certain in case she is level-1 or level-3, and that she chooses $B_{(2)}$, $A_{(3)}$, and $A_{(4)}$ with probabilities $1/2$, $1/6$, and $1/3$, respectively, in case she is level-2 or level-4. The maximum-likelihood estimate for the coordination game yields 44% of level-0 and 56% of levels 1 and above, while the estimate for the discoordination game

[33]Note that for the coordination game, Lk_{MOD}-POSTCOORD-INDL0 makes the same predictions as a Lk-POSTCOORD-INDL0 model. For the discoordination game, the two differ only slightly (and not at all in terms of the estimated level-k distributions).

yields 91% level-0, a combined 1% of level-1 and level-3 players and a combined 8% of level-2 and level-4 players.

Main Result 4. Basing the level-k model on individual post-game guessing-task estimates as $L0$ improves the model's fit to the coordination-game data. Yet, the maximum-likelihood estimate even of this model does not produce a prediction that would capture the essential features of the discoordination-game data.

The first part of Main Result 4 results from comparison of mean squared prediction errors in the last column of Table 8. The second part results from the fact that the predicted choice distribution also of the model Lk_{MOD}-POSTCOORD-INDL0 predicts the modal choice to be $B_{(2)}$ rather than $A_{(3)}$.[34]

5 Summary and discussion

The data gathered by Rubinstein et al. in their hide-and-seek experiments pose a serious challenge to Nash-equilibrium as a descriptive theory of behaviour. Up to today, two explanations have been proposed. Rubinstein and Tversky (1993; p. 402) claim that participants would fail to reason strategically and "[employ] a naïve strategy (avoiding the endpoints), that is not guided by valid strategic reasoning." If that were so, we might expect them to choose similarly to what they would pick if they had to click on any of the boxes without there being a game: according to the options' salience.[35] **Observation 4** establishes that salience-clicking is a bad predictor for aggregate behaviour in hide-and-seek games.

A second explanation has been proposed by Crawford and Iriberri (2007). They propose a level-k model that is based on salience and convert Rubinstein et al.'s account of what is salient into latent model parameters. They estimate the qualitative salience pattern to be such that "central A" is the least salient location, followed by "B", leaving the two "end As" as the most salient locations. Based on this salience pattern, Crawford and Iriberri (2007) present a model that fits the data almost as well as a benchmark model based on hard-wired payoff perturbations, and that outperforms any other model they study in terms of out-of-sample predictions. However, in all six salience-elicitation experiments I conducted, "B" turned out as the most salient location, while "central A" never was the (single) least salient location (**Observations 1** and **2**). The natural question to be

[34] χ^2-tests on the data under the hypothesis that the data stems from the model's predicted distribution yield $p \approx 0.21$ for the coordination game, and $p = 0.008$ for the discoordination game.

[35] Of course, "[in]valid strategic reasoning" might refer to other things apart from choice by salience. Yet, as long as we are not told more than the fact that participants avoid the endpoints, we are talking about a description of the data, rather than about a testable explanation for it.

answered was then whether a level-k model based on the empirically-elicited salience patterns would do equally well as a descriptor/predictor of behaviour. **Main Result 1** and **Observation 5** clearly show this is not the case, and that the best-fitting estimates, on top of having a poor fit to the data, exhibit level-k type distributions that are implausible. Coincidentally, the benchmark equilibrium models with salience-based payoff perturbations fit the data similarly badly when based on the empirically-elicited salience patterns (**Observation 7**).

Does this mean a level-k model cannot be used to account for behaviour in hide-and-seek games at all? The answer is no. By following an argument by Mehta et al. (1994) and modifying the model such that a level-k player will choose according to the salience of a location (consciously or not) whenever the player would be indifferent under pure payoff considerations, I obtain a model that has a fit of similar order as the best-performing models presented by Crawford and Iriberri (2007; **Main Result 2**). **Observation 6** lends empirical support to the idea behind the modification: when otherwise-indifferent players are assumed to mix uniformly, the model estimates show a large fraction of salience-guided level-0 behaviour; once these players mix according to salience, the estimated fraction of level-0 play is nill. The result is a better-fitting model that makes use of one parameter less.

Hargreaves Heap et al. (2014) have participants play a coordination and a discoordination game on the same action-set frames as the hide-and-seek game. Using this approach, they show at a general level that calibrating salience and the type distributions on hide-and-seek game data to predict behaviour in coordination and discoordination games is doomed to fail. However, to rebut the salience-based level-k model, they need to assume that the model applies to all three types of games. Crawford (2014) contests this assumption, arguing that team-reasoning rather than level-k should be applied to coordination games. Also, **Observations 3** and **8** suggest that players' salience perception may change to some (limited) degree across different games with the same action-set frames. In this light, the question of whether a salience-based level-k model can explain behaviour in different games remains unanswered. **Main Result 3** establishes that the conclusion Hargreaves Heap et al. draw is correct even if we restrict our focus to the discoordination game and account for a changing level-0: in this game, Crawford and Iriberri's preferred model predicts behaviour badly. When I subject the modified level-k model based on empirical salience measures to the same test, it does better, but it still clearly fails to explain the data.

Up to this point, all authors including myself have assumed that at least within each game, there is a unique level-0 that is the same for all players.[36] A

[36] A notable exception is Burchardi and Penczynski (2014), who allow for different guesses about the behaviour of non-strategically-acting players in a beauty-contest game.

closer look at the data from the experiments presented in this study reveals this assumption is unwarranted, too. Data from guessing-task experiments are by no means homogeneous in terms of what participants expect others to choose in a PICKING TASK. Therefore, we need to analyse behaviour taking into account heterogeneous salience perceptions. I do this exemplarily using the discoordination-game (and coordination-game) data. **Main Result 4** shows that basing a level-k model on individual post-game guessing-task estimates as level-0 can improve the model's fit to the data. Unfortunately, it also shows that a heterogeneous level-0 still does not allow to understand the data from the discoordination game.

Beyond a doubt, the results presented here pose a serious challenge to level-k theory. In this paper, I have been rather lenient with the theory, by allowing also higher-order salience to be a level-0 candidate as well as by adding a (plaubible) model modification that was able to accommodate the hide-and-seek data.[37] Nonetheless, even under these forgiving conditions, level-k cannot account for discoordination-game data even with a heterogeneous level-0. Unless level-k theory can be modified in a way that provides an explanation also for the results presented here, we should be hesitant to accept a level-k explanation for behaviour in hide-and-seek games. And yet, we need to bear in mind that there is no other model at hand that can explain the recurrent features of the hide-and-seek data. Also note that studies like Burchardi and Penczynski (2014) find empirical support for level-k-like reasoning in a clever design that allows to observe participants' *reasoning* rather than only their choices. What do the results mean, then? They may mean that only a subset of participants really follow level-k reasoning. For the remaining (majority of the) participants, we may have to look for different models to understand their behaviour.

Technical acknowledgements

All experiments were computerised using z-Tree (Fischbacher, 2007), participants were recruited using ORSEE (Greiner, 2004) and hroot (Bock, Baetge, and Nicklisch, 2014) with Mozilla Firefox. The statistical analyses were done using R (R Development Core Team 2001, 2012; Ihaka 1998) in combination with RKWard (Rödiger et al., 2012). All this was done on a computer running on KDE-based (KDE eV, 2012) Kubuntu, which required the use of wine for the programming of the experiments. The article was written using Kile.

[37] A recent working paper by Penczynski (2014) suggests we should be even more 'lenient' (or more realistic) and incorporate different level distributions for hiders and seekers. However, this does not help in explaining the data that seem to pose the most serious challenge, namely the discoordination-game data.

References

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

Burchardi, K. B. and Penczynski, S. P. (2014). Out of your mind: Estimating individual reasoning in one shot games. Games and Economic Behavior, 84:39–57.

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.

Christenfeld, N. (1995). Choices from identical options. Psychological Science, 6(1):50–55.

Costa-Gomes, M. A., Crawford, V. P., and Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. Econometrica, 69(5):1193–1235.

Crawford, V. P. (2014). A comment on 'How portable is level-0 behavior? A test of level-k theory in games with non-neutral frames' by Heap, Rojo-Arjona, and Sugden. Working paper.

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.

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.

Ho, T.-H., Camerer, C., and Weigelt, K. (1998). Iterated dominance and iterated best response in experimental 'p-beauty contests'. American Economic Review, 88(4):947–969.

Lewis, D. K. (1969). Convention: A Philosophical Study. Harvard University Press.

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.

Penczynski, S. P. (2014). Strategic thinking: The influence of the game. Working paper.

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. (1996). 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.

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

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.

Stahl, D. O. and Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. Games and Economic Behavior, 10(1):218–254.

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

von Neumann, J. (1953). A certain zero-sum two-person game equivalent to the optimal assignment problem. In Kuhn, H. W. and Tucker, A. W., editors, Contributions to the Theory of Games, volume II, pages 5–12. Princeton University Press.

Wolff, I. (2014). On the salience-based level-k model. Research Paper 94, Thurgau Institute of Economics.

Appendix A Hide-and-seek data

TABLE 1—AGGREGATE CHOICE FREQUENCIES IN RTH'S TREATMENTS

RTH-4	A	*B*	A	A
Hider (53; $p = 0.0026$)	9 percent	36 percent	40 percent	15 percent
Seeker (62; $p = 0.0003$)	13 percent	31 percent	45 percent	11 percent
RT-AABA–Treasure	A	A	*B*	A
Hider (189; $p = 0.0096$)	22 percent	35 percent	19 percent	25 percent
Seeker (85; $p = $ 9E-07)	13 percent	51 percent	21 percent	15 percent
RT-AABA–Mine	A	A	*B*	A
Hider (132; $p = 0.0012$)	24 percent	39 percent	18 percent	18 percent
Seeker (73; $p = 0.0523$)	29 percent	36 percent	14 percent	22 percent
RT-1234–Treasure	1	*2*	3	4
Hider (187; $p = 0.0036$)	25 percent	22 percent	36 percent	18 percent
Seeker (84; $p = $ 3E–05)	20 percent	18 percent	48 percent	14 percent
RT-1234–Mine	1	*2*	3	4
Hider (133; $p = $ 6E-06)	18 percent	20 percent	44 percent	17 percent
Seeker (72; $p = 0.149$)	19 percent	25 percent	36 percent	19 percent
R-ABAA	A	*B*	A	A
Hider (50; $p = 0.0186$)	16 percent	18 percent	44 percent	22 percent
Seeker (64; $p = $ 9E-07)	16 percent	19 percent	54 percent	11 percent

Notes: Sample sizes and p-values for significant differences from equilibrium in parentheses; salient labels in italics; order of presentation of locations to subjects as shown.

Table A.1: Table 1 from Crawford and Iriberri (2007).

Heinrich and Wolff (2012)	A	B	A	A
Hider (208)	15 percent	29 percent	30 percent	25 percent
Seeker (141)	9 percent	22 percent	51 percent	18 percent

Table A.2: Hide-and-Seek game data from Heinrich and Wolff (2012)

Appendix B Translated instructions

A Picking Task

Please choose one of the following four boxes and click on it!

□A □B □A □A

B Guessing Task [Post-Game Guessing]

[Before this experiment, w]e have asked 405 students at the University of Konstanz in a questionnaire to choose one out of four boxes that were [also] marked as follows: A, B, A, A. [The 405 students did not know anything about you or about the game that you just played.]

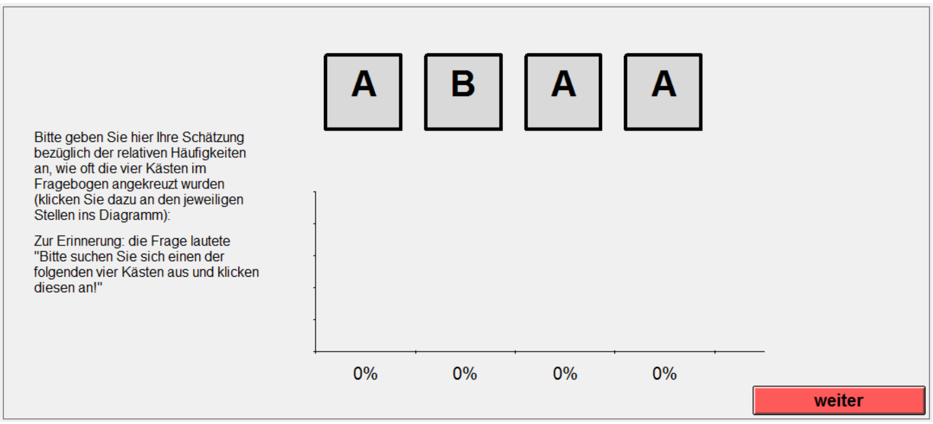
It is now your task to estimate as exactly as possible, what percentage of the students has chosen the respective boxes.

The closer your estimate is to the data we gathered, the more you can earn in this part of the experiment.

Details: in case your estimate does not deviate for any of the boxes by more than 5% from the true value, you receive 10 points; if the estimate deviates for at least one box by more than 5%, but for none by more than 10%, you receive 5 points; if the estimate deviates for at least one box by more than 10%, but for none by more than 20%, you receive 2 points, and otherwise you receive no points at all.[38]

Please enter here your estimate with respect to the relative frequencies of how often the four boxes were ticked in the questionnaire (to do so, click on the diagramme at the respective spots):

As a reminder, the question was: "please choose one of the following four boxes and click on it!"



[38] 10 points were equal to 0.50 Euros.

C Beauty Contest

Another beauty contest: [this classroom experiment was conducted right after discussing the $p-$beauty contest and level-k theory]

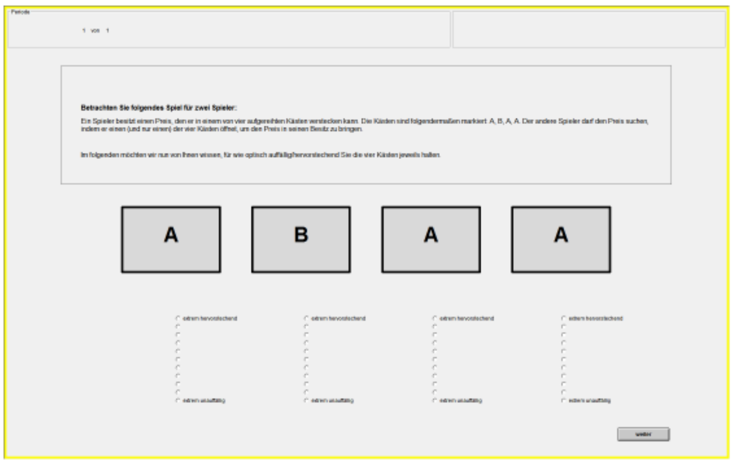
The question: which is the most salient box, which are the second, third, and fourth most salient boxes?
Among those stating the modal ordering, a prize of 12 Euros will be raffled off.
Use A_1, B_2, A_3, and A_4 to indicate the options.

D Rating Task [Post-Story Rating]

[Post-Story Rating: Consider the following game for two players:

One player owns a prize that he can hide in one of four aligned boxes. The boxes are marked as follows: A, B, A, A. The other player can search for the prize by opening one (and only one) of the four boxes, to take possession of the prize.]

In the following, we would like to know from you how optically salient/conspicuous you find the [Post-Story Rating: respective] four boxes [Rating Task: depicted below].



Appendix C Alternative payoff-perturbed game

Here, I present the game with hard-wired payoff perturbations when $A_{(4)}$ is least salient without $A_{(1)}$. Under the additional simplifying assumption that $A_{(1)}$ and $A_{(3)}$ are equally salient, we obtain the game shown in Table C.3.[39] Again, we should expect $e > 0$ and $f > 0$.

		Seeker			
		A$_{(1)}$	**B**$_{(2)}$	**A**$_{(3)}$	**A**$_{(4)}$
Hider	**A**$_{(1)}$	1 0	$0+f$ 1	0 1	$0-e$ 1
	B$_{(2)}$	0 $1-f$	$1+f$ $0-f$	0 $1-f$	$0-e$ $1-f$
	A$_{(3)}$	0 1	$0+f$ 1	1 0	$0-e$ 1
	A$_{(4)}$	0 $1+e$	$0+f$ $1+e$	0 $1+e$	$1-e$ $0+e$

Table C.3: The hide-and-seek game with payoff perturbations when $A_{(4)}$ is the single least salient location and $A_{(1)}$ and $A_{(3)}$ are equally salient.

Appendix D Full version of estimation-result Tables 5 and 6

On the next page, I include the Table-5 equivalent containing all estimated models, as well as the corresponding level-k distributions, including those reported in Table 6.

[39] This assumption can be based on the observations from the GUESSING TASK and, arguably, from the locations' average ranks in the BEAUTY CONTEST.

Specification	Data	$L0$	$L1$	$L2$	$L3$	$L4$	logL	MSE
Choices follow salience								
naïve-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$	rth	–	–	–	–	–	-1724	0.01271
naïve-postH&S-$B_{(2)}A_{(1)}\left[A_{(3)}A_{(4)}\right]$	rth	–	–	–	–	–	-1687	0.01647
naïve-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	rth	–	–	–	–	–	-1663	0.01226
Equilibrium models								
eqm$_0$	rth	–	–	–	–	–	-1641†	0.00967†
eqm$_+$-hyp-$\left[A_{(1)}A_{(4)}\right]$ ($e_H = 0.29, f_H = 0.25, e_S = 0.15, f_S = 0.15$)	rth	–	–	–	–	–	-1562†	0.00006†
eqm$_+$-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ ($e_H = 0.00, f_H = 0.06, e_S = 0.00, f_S = 0.05$)	rth	–	–	–	–	–	-1636	0.00909
eqm$_+$-hyp-$\left[A_{(1)}A_{(3)}\right]$ ($e_H = 0.08, f_H = 0.08, e_S = 0.17, f_S = 0.12$)	rth	–	–	–	–	–	-1608	0.00744
eqm$_+$-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$ / -rate-$B_{(2)}\left[A_{(1)}A_{(3)}A_{(4)}\right]$ ($e_H = 0.00, f_H = 0.06, e_S = 0.00, f_S = 0.05$)	rth	–	–	–	–	–	-1636	0.00909
CI's preferred model								
Lk-hyp-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$	rth	0.00	0.19†	0.32†	0.24†	0.25†	-1564†	0.00027†
Lk-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$	rth	0.38	0.14	0.00	0.48	0.00	-1616	0.00683
Lk-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	rth	0.09	0.28	0.01	0.62	0.00	-1664	0.01213
L**k**-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	rth	0.56	0.00	0.44	0.00	0.00	-1629	0.00830
CI's model with asymmetric $L0^{\S}$								
Lk-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$-asym	rth	–	0.00	0.15	0.64	0.21	-1632	0.00782
Lk-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$-asym	rth	–	0.08	0.26	0.32	0.34	-1603	0.00556
Lk-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$-asym	rth	–	0.58	0.21	0.00	0.21	-1636	0.00909
CI's model with salience-avoiding $L0^{\S}$								
Lk-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$-avoid	rth	–	0.00	0.79	0.06	0.15	-1632	0.00782
Lk-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$-avoid	rth	–	0.12	0.46	0.20	0.22	-1603	0.00556
Lk-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$-avoid	rth	–	0.21	0.00	0.21	0.58	-1636	0.00909
Modified level-k model								
Lk_{MOD}-neut-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$	rth	0.48	0.12	0.14	0.26	0.00	-1597	0.00310
Lk_{MOD}-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	rth	0.00	0.20	0.37	0.43	0.00	-1570	0.00097
Lk_{MOD}-postRate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	rth	0.22	0.00	0.65	0.01	0.12	-1621	0.00734
Choices follow salience								
naïve-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$	hw	–	–	–	–	–	-521	0.01654
naïve-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	hw	–	–	–	–	–	-487	0.01662
naïve-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	hw	–	–	–	–	–	-486	0.01525
Equilibrium models								
eqm$_0$	hw	–	–	–	–	–	-484	0.01436
eqm$_+$-hyp-$\left[A_{(1)}A_{(4)}\right]$ ($e_H = 0.38, f_H = 0.29, e_S = 0.10, f_S = 0.02$)	hw	–	–	–	–	–	-456	0.00109
eqm$_+$-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$ ($e_H = 0.00, f_H = 0.04, e_S = 0.00, f_S = -0.05$)	hw	–	–	–	–	–	-483	0.01467
eqm$_+$-hyp-$\left[A_{(1)}A_{(3)}\right]$ ($e_H = 0.12, f_H = 0.08, e_S = -0.03, f_S = -0.06$)	hw	–	–	–	–	–	-480	0.01712
eqm$_+$-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$ / -rate-$B_{(2)}\left[A_{(1)}A_{(3)}A_{(4)}\right]$ ($e_H = 0.00, f_H = 0.04, e_S = -0.03, f_S = -0.06$)	hw	–	–	–	–	–	-482	0.01485
CI's preferred model								
Lk-hyp-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$	hw	0.00	0.12	0.37	0.29	0.22	-456	0.00110
Lk-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$	hw	0.59	0.14	0.08	0.19	0.00	-476	0.01192
Lk-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	hw	0.70	0.07	0.00	0.23	0.00	-485	0.01496
L**k**-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	hw	0.61	0.00	0.33	0.06	0.00	-480	0.01259
CI's model with asymmetric $L0^{\S}$								
Lk-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$-asym	hw	–	0.08	0.10	0.65	0.17	-484	0.01320
Lk-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$-asym	hw	–	0.19	0.09	0.50	0.23	-465	0.00640
Lk-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$-asym	hw	–	0.49	0.29	0.00	0.22	-483	0.01367
CI's model with salience-avoiding $L0^{\S}$								
Lk-neut-$B_{(2)}\left[A_{(1)}A_{(3)}\right]A_{(4)}$-avoid	hw	–	0.00	0.75	0.07	0.18	-484	0.01320
Lk-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$-avoid	hw	–	0.03	0.56	0.17	0.24	-465	0.00639
Lk-rate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$-avoid	hw	–	0.29	0.00	0.22	0.49	-483	0.01367
Modified level-k model								
Lk_{MOD}-neut-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$	hw	0.30	0.18	0.14	0.38	0.00	-457	0.00152
best 'hump-shaped' type distribution	hw	0.26	0.18	0.18	0.38	0.00	-458	0.00166
Lk_{MOD}-postH&S-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	hw	0.00	0.23	0.28	0.49	0.00	-458	0.00143
Lk_{MOD}-postRate-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]^{\ddagger}$	hw	0.55	0.00	0.37	0.07	0.01	-477	0.01150

†indicates the estimate is taken from ci's paper. ‡ Better-performing specification from Rating Task and Post-Story Rating. $^{\S}\pi_0 \equiv 0$ imposed.

Table D.4: Full version of Tables 5 and 6 combined: maximum-likelihood estimates, log-likelihoods, and mean squared errors of the fit for the different models, using the data from Rubinstein, Tversky, and Heller's collected studies ("rth"), reproduced in Table 3 of Crawford and Iriberri (2007; "ci"). The table's lower half replicates the upper-half findings using Heinrich and Wolff's (2012; "hw") data.

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch