Erschienen in: Economics Letters ; 141 (2016). - S. 134-137
https://dx.doi.org/10.1016/j.econlet.2016.02.011

Elicited salience and salience-based level-*k*

Irenaeus Wolff

Thurgau Institute of Economics (TWI)/University of Konstanz, Hauptstrasse 90, 8280 Kreuzlingen, Switzerland

HIGHLIGHTS

- I measure salience in nine different ways.
- The elicited salience patterns are surprisingly similar.
- All measures differ clearly from the level-0 pattern needed to fit Hide&Seek data.
- Conversely, level-*k* based on any of the elicited patterns yields a poor data fit.

ABSTRACT

A level-*k* model based on a specific salience-pattern is the only model in the literature that accounts for behaviour in hide-and-seek games. This paper presents nine different experiments designed to measure salience. The elicited salience patterns tend to be similar, but none of them is similar to the pattern needed to allow the level-*k* model explain the hide-and-seek data. When based on any of the empirical salience measures, the salience-based level-*k* model does not fit the data well.

1. Introduction

Behaviour in many studies does not correspond to a Nash-equilibrium, in particular in one-shot games and early rounds of repeated games (Crawford et al., 2013). One of the main contestant models to account for unexperienced behaviour is the level-*k* model. In fact, it is the only model that has been shown to be able to account for behaviour in the hide-and-seek games presented in Rubinstein and Tversky (1993), Rubinstein et al. (1997), and Rubinstein (1999). In the archetype version of this 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 the seeker chooses the same box as the hider, the seeker gains the treasure, otherwise the hider keeps it. The typical choice distribution from experiments on the game differs markedly from the unique Nash prediction, uniform mixing by both players. The typical data set 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).[1]

Crawford and Iriberri (2007, henceforth CI) show that a level-*k* model anchored on a salience-seeking level-0 accounts for the observed data once the model is based on a specific salience-pattern. In this pattern, the "end As" are the most and "central A" the least salient locations. Hargreaves Heap et al. (2014) test the salience-based level-*k* model on a more general level and show that it cannot account simultaneously for data from hide-and-seek games, coordination games, and discoordination games if we assume that level-0 is the same for all games played on the same action-set frame. While Hargreaves Heap et al.'s paper casts doubt on the generalisability of the model, it does not address whether the model is a plausible explanation for the hide-and-seek data. To test whether this is the case, I measure salience in nine different ways, base CI's level-*k* model on the measured salience-patterns, and test whether any of the resulting models allows to explain the hide-and-seek data.

E-mail address: wolff@twi-kreuzlingen.ch.

[1] The experimental data of Rubinstein and co-authors are reported in Appendix A.

2. The salience-elicitation experiments

I examine nine experimental measures of salience. The point of this exercise is not to compare the different measures. The point is to test whether *any* of the measures yields a salience pattern that, being used as level-0 in CI's level-k model, would allow that model to account for the hide-and-seek game data.[2]

The first three experiments are measures of primary, secondary, and infinity-order salience, keeping the game description out. The fourth-to-sixth measures use the *secondary-salience* measure to explore the effect of introducing the game story (and whether an asymmetry follows from different player roles). Measures seven and eight provide alternative measures of *primary salience* with and without the hide-and-seek story, and measure nine is an alternative measure of *secondary salience*. To be precise, I look at the following experiments[3]:

PICKING TASK. Asking people to choose one of four boxes labelled "A", "B", "A", and "A" (Bardsley et al.,'s 2010 measure of *primary salience*), on a separate page of a post-experimental questionnaire (containing mostly items from the 16PF personality inventory) after an unrelated experiment. As a crucial complementary measure, I record response times for this task.

GUESSING TASK. Asking participants to estimate the relative click frequencies from the PICKING TASK (Bardsley et al.,'s 2010 measure of *secondary salience*).[4]

BEAUTY CONTEST. A beauty contest anchored in the question "which is the most salient box, which are the second, third, and fourth most salient boxes?" Conducted as a classroom experiment in the Experimental Methods course.[5]

POST-H&S GUESSING. A GUESSING TASK *after* participants had played the hide-and-seek game but before they got any feedback.[6]

POSTCOORD GUESSING. A GUESSING TASK after participants had played a coordination game on the A-B-A-A frame.[6]

POSTDISCOORD GUESSING. A GUESSING TASK after participants had played a discoordination game on the A-B-A-A frame.[6]

RATING TASK. Asking participants to rate the salience of each of the four boxes on an 11-point Likert scale ("extremely inconspicuous" to "extremely conspicuous").

POST-STORY RATING. A RATING TASK conducted after explaining the hide-and-seek game in a role-neutral format.[7]

POST-STORY RATE-GUESSING. A GUESSING TASK on the average POST-STORY RATING conducted after explaining the hide-and-seek game in a role-neutral format.[6,7]

None of the participants participated in more than one of the nine experiments.[8]

3. Results of the experiments

The results of the nine 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 $A_{(4)}$ is not more salient than $A_{(3)}$. Hence, Crawford and Iriberri's salience-seeking level-0 is mis-specified in terms of both the most salient and the least salient location.

For the first part, look at the second data column in Table 1. Treating the different salience measures as independent realisations of an underlying 'true' salience pattern and assuming that the next-salient candidate has an equal chance of being recorded as the most salient alternative in each of the eleven measures, we can compute the according binomial test's p-value to be $p = 1/2048$. Analogously, $A_{(3)}$ is more salient than $A_{(4)}$, with the same level of significance.

Observation 2. From the nine 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, $B_{(2)}[A_{(3)}A_{(1)}]A_{(4)}$, is observed in the GUESSING TASK, POSTCOORD GUESSING, the BEAUTY CONTEST, POST-STORY RATE-GUESSING, and possibly the RATING TASK.[9] The second pattern, $B_{(2)}A_{(3)}[A_{(1)}A_{(4)}]$, is observed in POST-H&S GUESSING, POSTDISCOORD GUESSING and possibly in POST-STORY RATING, while both the RATING TASK and the POST-STORY RATING patterns can be interpreted as $B_{(2)}[A_{(3)}A_{(1)}A_{(4)}]$.[10]

4. Model fit under the elicited salience patterns

I use the three elicited salience patterns as level-0 of a level-k model (instead of CI's $\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$), and repeat CI's maximum-likelihood estimation for the three new models on the same data they used.[11] Additionally, I estimated three 'hybrid' models in which level-0 follows *primary salience* and higher k-levels are determined by the corresponding *secondary-salience* measure. Table 2 presents the results. The hybrid model building on the *Picking-Task* data coincides with Lk-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$ in Table 2, while the model building on the RATING TASK performs insubstantially worse than the reported model Lk-HYB-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$.[12] To respond to the objection that salience may be influenced by culture but that the salience measures were obtained in a different country than the hide-and-seek data, I also include the model estimates for the hide-and-seek data from Heinrich and Wolff (2012).[13]

2 Hence, no care was taken to have similar numbers of observations in all experiments.

3 A translated version of the instructions to each task is provided in Appendix C.

4 If no frequency differed from the true value by more than 5% (10%/20%), participants earned 50 (25/10) Euro cents; the first of several (unknown) tasks participants faced in the experiment.

5 Amongst those stating the modal ordering, a prize of 12 Euros (USD 15.60) was raffled off.

6 Incentives as in the GUESSING TASK.

7 Participants did not play the game itself.

8 We used z-Tree (Fischbacher, 2007), ORSEE (Greiner, 2015), and hroot (Bock et al., 2014).

9 I subsume the POST-STORY RATE-GUESSING pattern here because Wilcoxon matched-pairs signed-ranks test clearly point to a difference between $A_{(3)}$ and $A_{(4)}$ ($p = 0.003$) but less clearly to one between $A_{(1)}$ and $A_{(3)}$ ($p = 0.097$).

10 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. None of the conclusions in this paper would change if we included $[B_{(2)}A_{(3)}]A_{(1)}A_{(4)}$ or $B_{(2)}A_{(1)}A_{(3)}A_{(4)}$ (from POST-STORY RATE-GUESSING) in the list of salience patterns.

11 The only (merely technical) difference is that I found the maximum-likelihood estimates by performing a complete grid search over all possible type-distributions (at the percent level) rather than using an algorithm.

12 I thank an anonymous referee for suggesting the 'hybrid' models.

13 For comparability, I include only the data obtained under the original instructions. Compared to the data CI use, this data has the additional advantage that it was obtained exclusively in the ABAA-frame, so that no further assumptions are needed of how to translate salience patterns from other settings, such as from the 1234-frame.

Table 1
Salience assessments of the four boxes denoted by “A”, “B”, “A”, and “A”.

	$A_{(1)}$	$B_{(2)}$	$A_{(3)}$	$A_{(4)}$
PICKING TASK (405 participants)				
relative click frequencies (%)	21	38	35	6
mean response times (seconds)	8.8	7.7	8.5	11.9
GUESSING TASK (72 participants)				
average estimated relative click frequency	21	41	22	15
BEAUTY CONTEST (30 participants)				
mean rank in beauty contest	2.3	1.5	2.5	3.6
POST-H&S GUESSING (156 participants)				
average estimated relative click frequency				
...by hiders (78 obs.)	19	38	24	19
...by seekers (78 obs.)	19	40	25	17
POSTCOORD GUESSING (72 participants)				
average estimated relative click frequency	19	50	18	14
POSTDISCOORD GUESSING (72 participants)				
average estimated relative click frequency	20	37	24	19
RATING TASK (90 participants)				
average conspicuousness reported (scale: 0–10)	5.7	7.5	5.6	5.3
POST-STORY RATING (90 participants)				
average conspicuousness reported (scale: 0–10)	3.8	7.4	4.3	4.0
POST-STORY RATE-GUESSING (84 participants)				
average estimated rating (scale: 0–10)	3.9	7.5	3.3	2.7

Table 2
Log-likelihoods and mean squared errors of the maximum-likelihood estimates of the indicated models. “RTH” refers to Rubinstein, Tversky, and Heller’s collected studies, whose data is reproduced in Table 3 of Crawford and Iriberri (2007). “HW” refers to Heinrich and Wolff (2012). The data from both studies is provided in Appendix A.

	RTH’s data		**HW’s data**	
Specification	**logL**	**MSE**	**logL**	**MSE**
Perfect fit	−1562[a]	0.00000	−452	0.00000
Lk-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$	−1564[a]	0.00027[a]	−456	0.00109
Lk-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$	−1616	0.00683	−476	0.01192
Lk-$B_{(2)}A_{(3)}\left[A_{(1)}A_{(4)}\right]$	−1635	0.00854	−485	0.01514
Lk-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$[b]	−1629	0.00830	−480	0.01259
Lk-HYB-$B_{(2)}\left[A_{(3)}A_{(1)}A_{(4)}\right]$[b]	−1635	0.00903	−482	0.01349
Mixed-strategy equilibrium	−1641[a]	0.00967[a]	−484	0.01436

[a] Indicates the estimate is taken from CI’s paper.
[b] The better-performing specification from RATING TASK and POST-STORY RATING.

Observation 3. The measured-salience-based estimates for CI’s model do not fit the data, and outperfrom the mixed-strategy Nash prediction only unsubstantially.

Observation 3 can be verified by a look at the mean squared errors in the right-most column in Table 2, comparing the specification Lk-$B_{(2)}\left[A_{(3)}A_{(1)}\right]A_{(4)}$ to specification Lk-$\left[A_{(1)}A_{(4)}\right]B_{(2)}A_{(3)}$ and to that of the mixed-strategy equilibrium. The result is even stronger for RTH’s data, in the central column of Table 2. Contrary to what we should expect, none of the estimated level-k distributions is hump-shaped.[14]

5. Discussion

The results presented here pose another serious challenge to the salience-based level-k model. This is despite the fact that I have been rather lenient with the theory, by allowing also higher-order salience to be a level-0 candidate. Nonetheless, even under these forgiving conditions, the salience-based level-k model cannot account even for the hide-and-seek game data it was constructed for. Unfortunately, this means we are left without an explanation for the data other than Rubinstein et al.’s (1997) – unsatisfactory – account of participants choosing “a naïve strategy (avoiding the endpoints), that is not guided by valid strategic reasoning”.

[14] See Table B.3 in Appendix B.

Acknowledgments

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. Further, I thank Vincent Crawford, Martin Dufwenberg, Urs Fischbacher, Shaun Hargreaves Heap, David Rojo Arjona, Timothy Salmon, Dirk Sliwka, Robert Sugden, Marie-Claire Villeval, Roberto Weber, the lively research group at the Thurgau Institute of Economics (TWI), participants of the 2013 GfeW Meeting and the 2014 ESA European Meeting, as well as a number of anonymous referees for helpful comments. Financial support by the University of Konstanz’ Young Scholar Fund (no. 1414/54740/83/83954913) is gratefully acknowledged.

References

Bardsley, N., Mehta, J., Starmer, C., Sugden, R., 2010. Explaining focal points: Cognitive hierarchy theory versus team reasoning. Econ. J. 120, 40–79.
Bock, O., Baetge, I., Nicklisch, A., 2014. hroot: Hamburg registration and organization online tool. Eur. Econ. Rev. 71, 117–120.
Crawford, V.P., Costa-Gomes, M.A., Iriberri, N., 2013. Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. J. Econ. Lit. 51 (1), 5–62.
Crawford, V.P., Iriberri, N., 2007. Fatal attraction: Salience, naïveté, and sophistication in experimental ‘hide-and-seek’ games. Am. Econ. Rev. 97 (5), 1731–1750.

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

Greiner, B., 2015. An online recruitment system for economic experiments. J. Econ. Sci. Assoc. 1 (1), 114–125.

Hargreaves Heap, S., Rojo Arjona, D., 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., Wolff, I., 2012. Strategic reasoning in hide-and-seek games: A note. Research Paper 74. Thurgau Institute of Economics.

Rubinstein, A., 1999. Experience from a course in game theory: Pre- and postclass problem sets as a didactic device. Games Econom. Behav. 28 (1), 155–170.

Rubinstein, A., Tversky, A., 1993. Naive strategies in zero-sum games. Working Paper 17–93. The Sackler Institute of Economic Studies.

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