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

Vincent P. Crawford[1]

10 July 2014

Abstract: Rubinstein and Tversky (1993), Rubinstein, Tversky, and Heller (1996), and Rubinstein (1999) reported experiments that elicited initial responses to coordination, discoordination, and hide and seek games with non-neutral decision labelings, in which behavior responded to labeling. Crawford and Iriberri (2007ab) proposed a level-*k* model to account for the results for hide and seek games. Heap, Rojo-Arjona, and Sugden (2014) report an experimental test of Crawford and Iriberri's model, concluding that it lacks portability. This comment seeks to clarify Heap et al.'s interpretation of their results and their account of Crawford and Iriberri's analysis, and to better identify the way forward.

Keywords: behavioral game theory, experimental game theory, strategic thinking, level-*k* models, coordination, salience

[1] Department of Economics, University of Oxford, Manor Road Building, Oxford OX1 3UQ and All Souls College, Oxford OX1 4AL, UK, e-mail vincent.crawford@economics.ox.ac.uk; and Department of Economics, University of California, San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0508, USA. I thank Miguel Costa-Gomes and Nagore Iriberri for their extremely helpful advice. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 339179. The contents reflect only the author's views and not the views of the ERC or the European Commission, and the European Union is not liable for any use that may be made of the information contained therein. The University of Oxford and All Souls College also provided research support.

1. Introduction

This paper is a comment on Heap, Rojo-Arjona, and Sugden's (2014; "HRS") experimental test of the portability of Crawford and Iriberri's (2007ab; "CI") level-*k* model of Rubinstein and Tversky's (1993), Rubinstein, Tversky, and Heller's (1996), and Rubinstein's (1999) ("RTH") experimental results for hide-and-seek games with non-neutral decision labelings. My goals are to clarify HRS's interpretation of their results and their account of CI's analysis, and to better identify the way forward.

I begin by summarizing the background for HRS's analysis. Traditional game theory assumes that if context such as how players' decisions are identified does not directly affect preferences, it has no effect on behavior. However, it has been known since Schelling (1960) that such identifications can affect behavior via the patterns of salience they create, even if the identifications are via abstract labels free of connotations that might affect preferences. This view is supported by RTH's (among many other) experiments, in which subjects' decisions in coordination, discoordination, and hide-and-seek games varied with decision labelings, some abstract and others with connotations.

CI studied RTH's hide-and-seek treatments with abstract labeling, as especially informative. RTH's hide-and-seek subjects deviated systematically from equilibrium, with both hiders and seekers tending to favor the decision whose position and labeling RTH (and CI) argued made it "least salient", and with seekers favoring that decision even more than hiders.[2] HRS call this pattern the "fatal attraction" pattern, after CI's title. It is surprising to observe systematic deviations from equilibrium in a game where the equilibrium strategy is so obvious and its rationale so strong. The fatal attraction pattern is doubly surprising because equilibrium and noisy generalizations like quantal response equilibrium (McKelvey and Palfrey 1995; "QRE") preclude role-asymmetric patterns of deviation in RTH's hide-and-seek games, despite their asymmetric payoff structures.[3]

[2] In hide-and-seek games, any pure or mixed strategy is a best response to the equilibrium beliefs. But systematic deviations of aggregate choice frequencies from equilibrium mixed-strategy probabilities must with high probability have a cause that is partly common across players, and so are indicative of systematic deviations from equilibrium.

[3] In RTH's hide-and-seek games, QRE coincides with equilibrium for any symmetric noise distributions. With payoff perturbations added to reflect players' possible instinctive reactions to salience, both equilibrium and QRE can explain why central A is more prevalent, but QRE predicts that it is even more prevalent for hiders than seekers, the

To explain the pattern CI proposed a level-*k* model, in which players anchor their beliefs in a nonstrategic initial assessment of others' likely responses to the game called "*L0*" and then adjust them via iterated best responses, with *L1* best responding to *L0*, *L2* to *L1*, and so on. Players' levels are heterogeneous, drawn from a population distribution concentrated on the lowest levels but excluding *L0*, which most evidence suggests exists only in the minds of higher levels (Crawford, Costa-Gomes, and Iriberri 2013; "CCGI").

Even if *L0*'s population frequency is zero, its specification is of central importance. The usual uniform random specification is behaviorally implausible when labeling creates non-neutral patterns of salience, and would in any case make *Lk* decisions mimic equilibrium in hide-and-seek games. Instead CI allowed *L0* to favor decisions with salient labels, with higher levels responding to salience only through their iterated best responses to such an *L0*. To avoid begging the question of the fatal attraction pattern's asymmetry in hiders' and seekers' behavior, CI constrained *L0* and the population level frequencies to be the same in both player roles.

Unlike equilibrium and QRE, CI's level-*k* model predicts role-asymmetric decision distributions in response to the asymmetric payoff structure of hide-and-seek games. With a level distribution estimated for RTH's data, CI's model fits better than the alternatives, and suggests a mechanism to explain the fatal attraction pattern.[4] (CI, p. 1738, explain the mechanism, which depends on the heterogeneity of levels' decisions.)

Aware that their estimation gives their level-*k* model more freedom than equilibrium or QRE, CI tested the model's portability by adapting it to the different payoff structures and salience patterns of O'Neill's (1987) and Rapoport and Boebel's (1992) hide-and-seek-like games, using the level frequencies estimated from RTH's data and defining *L0* via the same principles as for RTH's games.[5] The adapted model describes O'Neill's subjects' early responses well, and Rapoport and Boebel's subjects' fairly well.

---

opposite of what RTH observed (CI 2007b). At this writing the fatal attraction pattern is still inconsistent with any model of which I am aware, except CI's salience-sensitive level-*k* model discussed next.

[4] HRS (pp. 1133-1134) imply that CI's estimation of which kind of salience *L0* responds to more strongly makes CI's model unparsimonious, but this estimation adds only a single binary-valued parameter: the smallest possible non-zero price to pay for accommodating a non-zero response to salience.

[5] CI's adapted model was specified before they gained access to O'Neill's and Rapoport and Boebel's data. HRS's footnote 23 criticizes CI's portability analysis, saying "…in effect CI use a new *L0* specification for each of the

HRS's experiments test the portability of CI's model across games whose structures go well beyond hide-and-seek-like games.[6] HRS's premise is that a useful model must be well-defined in advance for any game, and should be evaluated via ex ante hypothesis testing, not ex post model fitting (HRS, p. 1135).[7] Like RTH's subjects, HRS's played a series of hide-and-seek, coordination, and discoordination games. Unlike RTH's subjects, HRS's played such games within subjects, in groups with varying structures but with the number and labeling of decisions constant within groups and across player roles.

HRS use their results to test a level-*k* model in the spirit of CI's, assuming that *L0* has zero frequency and higher levels respond to labeling only through their iterated best responses to a salience-sensitive *L0*, and constraining *L0* and the population level frequencies to be constant across player roles in a given game. HRS argue that because *L0* is often motivated as nonstrategic (e.g. in CCGI, p. 14, as "a strategically naïve initial assessment of others' likely responses to the game"), a level-*k* model should hold equally across hide-and-seek, coordination, and discoordination games, with *L0* constant across the games within each of HRS's groups with analogous decisions and labelings.

From this constant-*L0* assumption HRS derive cross-game restrictions on behavior. In each of their groups of games, the decision labelings make one label, the "oddity", uniquely different from the others. In their coordination games, the oddity is subjects'

---

games". In fact CI used plausible general principles of salience to adapt *L0* to the new hide-and-seek-like games with different patterns of salience. O'Neill's labeling, for instance, was A-2-3-J (playing cards), for which CI took A and J (only) to be salient, as both face cards and end locations as the cards were presented to subjects. Rapoport and Boebel's labeling was C-L-F-I-O, for which CI took C, F, and O (only) to be salient because of their locations, on the assumption that no labels are salient per se. These identifications stop short of the general theory of salience HRS seem to wish for, but their implications in these applications are hardly controversial, and I would be suspicious of any pure theory of salience claiming full generality across cultures.

[6] Another notable example of cross-game testing is Georganas, Healy, and Weber (2014), who study the portability of level-*k* models across "undercutting" matrix games and Costa-Gomes and Crawford's (2006) two-person guessing games, finding that individual subjects' estimated levels are not very consistent across types of game, and that the aggregate level distribution is (p. 3) "remarkably stable across undercutting games, but quite unstable across two-person guessing games." In a notable example of cross-*role* testing, Penczynski (2014) contrasts hiders' and seekers' decisions in a design whose player roles are filled by two-person teams, whose communications are monitored to gain further insight into their thinking. He finds significantly higher levels for seekers than hiders. Burchardi and Penczynski (2014) use similar methods to identify *L0* in level-*k* models for *n*-person guessing games. CGCI (pp. 24-25) discuss these papers in much more detail.

[7] I agree that ex ante hypothesis testing is preferable for model evaluation, other things equal. But it is seldom the most efficient way to discover and interpret new facts about behavior. For that and other reasons, ex post model fitting also has a long and honorable tradition in empirical and experimental economics.

modal and even majority choice for almost all groups/labelings. As those games are symmetric across player roles, in a level-*k* model the oddity must be *L0*'s modal choice.

With *L0* constant across games and player roles within a group, the oddity's average frequency across discoordinators, hiders, and seekers must then be disproportionately high (HRS, p. 1138, Implication 1). However, HRS's subjects chose the oddity far less often in discoordination and hide-and-seek games than in coordination games, to the point where its frequency is usually too low to be consistent with their constant-*L0* level-*k* model (HRS, Table III). HRS conclude that "…it would seem hard to be optimistic about finding…a general theory of *L0* behavior" (p. 1135); and by implication that level-*k* models lack the portability needed to be useful.

Given the interest of modeling how labeling and context more generally affect strategic behavior, experiments like HRS's that test such models across games with a wider range of strategic structures are welcome. I believe, however, that HRS's interpretation of their results and their account of CI's analysis need clarification; and that the way forward can be better identified.

The rest of this paper is organized as follows. Section 2 argues that a level-*k* model, even with a nonstrategic *L0*, does *not* plausibly extend from HRS's discoordination and hide-and-seek games to their pure coordination games with analogous decisions and labeling. In pure coordination games like HRS's, both intuition and existing evidence point instead toward "team reasoning", whereby "each player chooses the decision rule which, if used by all players, would be optimal for each of them" (Bardsley, Mehta, Starmer, and Sugden 2009, p. 40).

Together with Section 3's arguments for the prevalence of level-*k* thinking in hide-and-seek games, the prevalence of team reasoning in pure coordination games suggests that the most promising route to a general model of strategic responses to context and labeling is not via the blanket rejection of level-*k* models that HRS seem to advocate (p. 1135); but rather via experiments in the service of constructive modeling, with the goal of identifying a robust hybrid of team reasoning, level-*k* thinking, and perhaps other kinds of thinking not yet codified, with a clear specification of which settings evoke which kinds of thinking.[8] As Crawford, Gneezy, and Rottenstreich (2008, p. 1448) said of their experimental results for (both pure and impure) coordination games, "Overall, our results suggest a synthesis of level-*k* thinking and team reasoning in which team reasoning supplements or supplants level-*k* thinking in some settings."

In Section 3 I respond to the doubts HRS (p. 1148) express about CI's identification of the fatal attraction pattern in RTH's hide-and-seek treatments with abstract labelings, explaining CI's rationale for their identification in more detail. Given HRS's doubts, they effectively ignore the fatal attraction pattern and the similar (though not identical; see footnote 10) pattern in their own data for hide-and-seek games with abstract labeling, offering no model or conjecture about why those subjects chose as they did. Whatever doubts remain, RTH's and HRS's hide-and-seek subjects' deviations from equilibrium suggest iterated-best-response reasoning, not fixed-point reasoning: an important clue that any general model will need a level-*k* as well as a team-reasoning component.

Section 4 reiterates the need for further experiments to identify a robust hybrid of team reasoning, level-*k* thinking, and other kinds of thinking, and to delineate which kinds of settings evoke which kinds of thinking. Progress along these lines will depend on developing a more general definition of team reasoning, which to my knowledge has not yet been clearly defined for "impure" coordination games whose players do not have identical preferences. Section 4 concludes by arguing that such experiments should avoid decision labelings with strong connotations like those in some of HRS's games, which sacrifice experimental control of preferences while gaining little of value in return.

2. Model Specification across Hide-and-Seek and Coordination Games

Recall that HRS's subjects played a series of hide-and-seek, coordination, and discoordination games within subjects, in groups with varying structures but with the number and labeling of decisions constant within groups and across player roles.[9] HRS

---

[8] HRS (p. 1135) acknowledge that a level-*k* model with a constant *L0* is not the only possibility: "Of course, we cannot claim that the portability property we test is implied by every possible general hypothesis about *L0* behavior." But they make no attempt to explore or discuss non-constant-*L0* alternatives; nor do they acknowledge that the most empirically promising alternative model for pure coordination games in the literature is not even a level-*k* model with a non-constant *L0*, but rather team reasoning.

[9] Payoffs were determined entirely by the relationship between the labels of players' decisions. In the coordination games players who choose decisions with the same label receive payoff one; otherwise both receive zero. In the

use their results to test a level-*k* model in the spirit of CI's, assuming that *L0* has zero frequency and higher levels respond to labeling only through their iterated best responses to a salience-sensitive *L0*, and constraining *L0* and the population level frequencies to be constant across player roles in a given game. HRS argue that because *L0* is often motivated as "nonstrategic" (e.g. CCGI, p. 14), while the differences across games within their groups are strategic, a level-*k* model should hold equally across their hide-and-seek, coordination, and discoordination games, with *L0* constant across the games within each of HRS's groups with analogous decisions and labelings.

In this section I argue, however, that a level-*k* model with a nonstrategic *L0* does *not* plausibly extend from HRS's discoordination and hide-and-seek games to their pure coordination games with analogous decisions and labeling. It is behaviorally far less plausible to assume that a strategically naïve person cannot distinguish pure coordination from zero-sum games at all, than to assume that s/he cannot model others' responses to incentives; and CI's and CCGI.'s motivation of *L0* was meant to convey only the latter.

Further, it is already well known that coordination games may trigger team reasoning, which is inherently different from level-*k* thinking in that it relies on fixed-point reasoning (Bardsley et al. 2009, p. 40). It is inappropriate to conflate fixed-point team reasoning rules and level-*k* rules based on iterated best responses to a given prior; and it should be no surprise that experimental tests reject such a conflation.

3. The Fatal Attraction Pattern

Recall that CI focused on RTH's hide-and-seek treatments with abstract labelings. If such labelings are free of connotations that might affect preferences, those treatments induce zero-sum two-person games with unique equilibria, in which players randomize uniformly without regard to labeling. Despite the strength of the normative justification for such equilibria, RTH's subjects' decision frequencies deviated systematically from

---

discoordination games players who choose decisions with different labels receive payoff one; otherwise both receive zero. In the hide-and-seek games, if players choose decisions with the same label, one player, the "seeker", receives payoff one and the other, the "hider", receives zero; and vice versa if their decisions have different labels.

equilibrium, in patterns that were sensitive to salience, and those patterns partly recurred in HRS's own data for hide-and-seek treatments with abstract labeling.[10]

The key to understanding those patterns is identifying the non-neutral salience landscapes subjects were responding to. CI (Section I) argued that all six of RTH's hide-and-seek treatments with abstract labeling had analogous landscapes and the same qualitative pattern of deviations from equilibrium, the so-called fatal attraction pattern. CI took the robustness of the pattern as a signal that something especially informative was happening in those treatments. By contrast, HRS (p. 1148) express doubts about CI's identification of the pattern. Given HRS's doubts, they effectively ignore the pattern and similar patterns in their own data for hide-and-seek treatments with abstract labelings, offering no model or conjecture about why subjects chose as they did in those treatments.

In this section I explain CI's rationale for their identification of the fatal attraction pattern, so readers can judge for themselves. Start with the treatment CI (Table I) called "RTH-4". In that treatment subjects' decisions were ordered left to right and labeled "A", "B", "A", "A". RTH and CI identified the "B" decision as salient via the uniqueness of its label; and the "end A" decisions as also salient, on which RTH cited Christenfeld (1995).[11] Given these saliencies, RTH and CI viewed "central A" as the "least salient" decision. In this salience landscape, the least salient decision was modal for both hiders and seekers, and even more frequent for seekers: the fatal attraction pattern.

CI argued that RTH's five other hide-and-seek treatments with abstract labeling had analogous salience landscapes, and that the fatal attraction pattern extends qualitatively to them. CI's arguments are based on RTH's intuitive conjectures about the least salient decisions in their various labelings; on Rubinstein and Tversky's (1993) observation that their "treasure" and "mine" treatments have the same normal form with player roles

---

[10] In HRS's treatments with games like those of RTH's that CI studied, the analogue of central A is still modal for seekers and even more prevalent for seekers than hiders. But for hiders the analogues of the end As are now 11% more frequent than that of central A, which is still chosen above chance (HRS, footnote 22). For the reasons explained in CI (2007b), this pattern remains a puzzle for models other than a salience-sensitive level-*k* model.

[11] In their footnote 21 HRS criticize CI's argument on this point: "The latter claim is supported by an unexplained citation of an experiment by Christenfeld (1995), which in fact found that when individuals pick from a row of identical items, they tend to avoid the end locations (CI, p. 1732)." As CI explained, this argument was first made by Rubinstein, Tversky, and Heller (1996, p. 401). Further, HRS's criticism rests on their implicit assumption that Christenfeld's subjects (in a decision problem with no clear rewards) wished to favor salient decisions rather than avoid them. Rubinstein and Tversky (1993, p. 4) made clear that in their view, the "end A" locations are salient.

interchanged and, mutatis mutandis, evoked roughly the same responses; and on the assumption that behavior is determined by the normal rather than extensive form.[12]

With regard to RTH's conjectures about the least salient decisions, no one doubts that uniquely labeled decisions are salient; and RTH's belief that end locations are salient is at least plausible (footnote 11). Rubinstein and Tversky's identification of "3" as least salient in their 1-2-3-4 treatments is less clear than that "central A" is least salient in A-A-B-A or A-B-A-A treatments because based on position alone, but it too is plausible.[13]

CI's assumption that behavior is determined by the normal rather than the extensive form extends a common assumption in equilibrium-based game theory (e.g. Kohlberg and Mertens 1986) to nonequilibrium models, in a way that is consistent with RTH's findings (footnote 12). To my knowledge no theory using information on the extensive form has yet been proposed for games like these.

Overall, CI's analogies use theoretical intuitions to group observations in the process of identifying a model to make sense of otherwise puzzling patterns, a long-standing and informative practice in empirical and experimental economies. Whether or not one agrees with every detail of CI's arguments identifying the fatal attraction pattern in RTH's data, subjects did deviate from equilibrium in role-asymmetric patterns that CI showed are indicative of some kind of iterated-best-response reasoning, not the fixed-point reasoning on which alternatives to level-*k* models are based. This is an important clue that a robust hybrid model of behavior across games like RTH's and HRS's will need a level-*k* as well as a team-reasoning component.

HRS go on to argue that, even if the fatal attraction pattern is real, level-*k* models are flexible enough to account for a large minority of all possible patterns, so that it is not especially surprising that CI's level-*k* model can account for them. Specifically, they note that there are only 18 possible qualitative patterns, and that some plausible form of a

[12] RTH introduced their mine treatments to test whether the difference in the extensive form due to the fact that hiders must hide before seekers seek, might explain subjects' deviations from equilibrium; and found that they did not.

[13] HRS's claim that the assumed least salience of "3" was "[w]ithout further explanation" ignores CI's (p. 1736) reference to Rubinstein and Tversky's admittedly vague conjecture. Even omitting this treatment, the others share a consistent pattern. HRS also critically mention CI's assumption that subjects choose the end A locations with equal frequencies, which is an expository simplification suggested by the data on which nothing important turns.

level-*k* model can account for seven of them. This argument by counting qualitative frequency patterns implicitly treats the patterns as random by giving them equal weight.

By contrast, CI argued that their model's explanation of the fatal attraction pattern is surprising against the highly *non*-random background of alternative theories, none of which but their salience-sensitive level-*k* model can explain it. Put another way, if the pattern is real, it is just the kind of surprising regularity with the power to discriminate among alternative models that empirically oriented economists usually seek.

4. The Way Forward

Section 3's and Section 4's arguments make clear that the best chance for a general model of strategic responses to context across games with a variety of structures like RTH's and HRS's is not the blanket rejection of level-*k* models HRS seem to advocate, but rather via experiments in the service of constructive modeling, with the goal of identifying a robust hybrid of team reasoning, level-*k* thinking, and perhaps other kinds of thinking, with a clear specification of which settings evoke which kinds of thinking.[14]

Progress along these lines will depend on developing a more general, theoretically motivated but empirically grounded definition of team reasoning, which to my knowledge has not yet been clearly defined for impure coordination games.

Although connotations are not the main issue in my critique of HRS's analysis, future experiments should avoid decision labelings with strong connotations like those in some of HRS's groups, some as emotionally loaded as "Hitler".[15] It is a commonplace in the marketing literature that labels with connotations influence subjects' choices, and they plainly did so in HRS's experiments.[16] Because such influences are not independently observable, they sacrifice control of preferences while gaining little in return.[17]

[14] The results of Bardsley et al. (2009) and Crawford et al. (2008) may provide a start. Bardsley et al. (2009) find cross-country variations in the occurrence of team reasoning. Crawford et al. (2008) find evidence of something in the spirit of team reasoning in some of their treatments with impure as well as pure coordination games.

[15] HRS (p. 1141) state that they used labels with connotations to "maintain subjects' interest and attention". I suggest that subjects' interest and attention are better maintained by making the experimental tasks engaging.

[16] Those influences would need to be modelled to draw useful inferences about theories of strategic behaviour. CI did not take a position on how to model the influence of connotations, because they focused on RTH's treatments with abstract labels. But CI (2007b) showed how to use payoff perturbations to model similar influences.

[17] Such influences might even change the strategic structure, undermining the cross-game implications on which HRS's tests depend. HRS (footnote 7) seek to address this criticism by allowing *L0* to respond to salience, while restricting higher levels to respond only through their iterated best responses to such an *L0*. Because they assume as

References

Bardsley, Nicholas, Judith Mehta, Chris Starmer, and Robert Sugden (2009): "Explaining Focal Points: Cognitive Hierarchy Theory *versus* Team Reasoning." *Economic Journal*, 120(2): 40-79.

Burchardi, Konrad B., and Stefan P. Penczynski (2014): "Out of Your Mind: Eliciting Individual Reasoning in One Shot Games." Games and Economic Behavior, 84 (1): 39–57.

Christenfeld, Nicholas (1995): "Choices from Identical Options." *Psychological Science*, 6(1): 50-55.

Costa-Gomes, Miguel A., and Vincent P. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review*, 96(5): 1737-1768.

Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri (2011): "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." *Journal of Economic Literature*, 51(1): 5-62.

Crawford, Vincent P., Uri Gneezy, and Yuval Rottenstreich (2008): "The Power of Focal Points is Limited: Even Minute Payoff Asymmetry May Yield Large Coordination Failures." *American Economic Review*, 98(4): 1443-1458.

Crawford, Vincent P., and Nagore Iriberri (2007a): "Fatal Attraction: Salience, Naivete, and Sophistication in Experimental Hide-and-Seek Games." *American Economic Review*, 97(5): 1731-1750.

Crawford, Vincent P., and Nagore Iriberri (2007b): Web appendix for "Fatal Attraction: Salience, Naivete, and Sophistication in Experimental Hide-and-Seek Games." http://www.aeaweb.org/aer/data/dec07/20050133_app.pdf .

Georganas, Sotiris, Paul J. Healy, and Roberto A. Weber (2014): "On the Persistence of Strategic Sophistication", unpublished.

CI did that *L0* players exist only in the minds of higher levels, this rules out any influence of connotations, unless connotations influence the results for their coordination games enough to change the discrete choice of *L1*. Although this does not happen in HRS's data, ruling out such responses prevents HRS from giving a full account of what their subjects are doing, much as would a version of consumer theory that rules out any effect of brands.

Heap, Shaun Hargreaves, David Rojo-Arjona, and Robert Sugden (2014): "How Portable is Level-0 Behavior? A Test of Level-k Theory in Games with Non-neutral Frames." *Econometrica*, 82(3): 1133–1151.

Kohlberg, Elon, and Jean-François Mertens (1986): "On the Strategic Stability of Equilibria." *Econometrica*, 54(5): 1003–1037.

McKelvey, Richard, and Thomas Palfrey (1995): "Quantal Response Equilibria for Normal-Form Games." *Games and Economic Behavior*, 10(1): 6-38.

O'Neill, Barry (1987): "Nonmetric Test of the Minimax Theory of Two-Person Zerosum Games." *Proceedings of the National Academy of Sciences of the United States of America*, 84(7): 2106-2109.

Penczynski, Stefan (2014): "Strategic Thinking: The Influence of the Game", unpublished.

Rapoport, Amnon, and Richard Boebel (1992): "Mixed Strategies in Strictly Competitive Games: A Further Test of the Minimax Hypothesis." *Games and Economic Behavior*, 4(2): 261-283.

Rubinstein, Ariel (1999): "Experience from a Course in Game Theory: Pre and Post-Class Problem Sets as a Didactic Device." *Games and Economic Behavior* 28(1): 155–170; "Second Edition" posted at http://arielrubinstein.tau.ac.il/99/gt100.html .

Rubinstein, Ariel, and Amos Tversky (1993): "Naïve Strategies in Zero-Sum Games." Working Paper 17-93, Sackler Institute of Economic Studies, Tel Aviv University.

Rubinstein, Ariel, Amos Tversky, and Dana Heller (1996): "Naïve Strategies in Competitive Games." In *Understanding Strategic Interaction: Essays in Honor of Reinhard Selten*, ed. Wulf Albers, Werner Güth, Peter Hammerstein, Benny Moldovanu, and Eric van Damme, 394–402. Berlin: Springer-Verlag.

Schelling, Thomas. 1960. *The Strategy of Conflict*. Cambridge: Harvard University Press.