Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 196 (2021) 105291

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Plasticity of strategic sophistication in interactive decision-making

Davide Marchiori [a], Sibilla Di Guida [b,*], Luca Polonio [c,d]

[a] Department of Marketing and Management, University of Southern Denmark, Denmark
[b] Department of Economics, University of Southern Denmark, Denmark
[c] Molecular Mind Laboratory (MoMiLab), IMT School for Advanced Studies Lucca, Italy
[d] Center for Mind/Brain Sciences – CIMeC, University of Trento, Italy

Received 21 November 2019; final version received 31 May 2021; accepted 3 June 2021
Available online 15 June 2021

Abstract

We propose an experimental eye-tracking study to test how strategic sophistication is shaped by experience in 3×3 two-person normal-form games. Although strategic sophistication has been shown to be linked to a variety of endogenous and exogenous factors, little is known about how it is affected by previous interactive decisions. We show that complete feedback in previous games can significantly enhance strategic sophistication, and that games that in principle provide equivalent learning opportunities lead instead to substantially different learning outcomes. Specifically, only repeated play with feedback of games that emphasize strategic interdependence significantly enhances strategic learning, producing an increase in the frequency of equilibrium play and a shift of attention to the incentives of the counterpart. Moreover, we find that the type of learning underlying newly gained strategic skills can vary substantially across players. Whereas some players eventually learn to visually analyze the payoff matrix consistently with equilibrium reasoning, others appear to use experience with previous interactions to devise simple heuristics of play. Our results have implications for theoretical and computational modeling of learning.

JEL classification: C72; C73; C91

Keywords: Learning; Strategic sophistication; Eye-tracking; Experimental economics; Repeated games

* Corresponding author at: Department of Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark.
E-mail address: sidg@sam.sdu.dk (S. Di Guida).

https://doi.org/10.1016/j.jet.2021.105291
0022-0531/

1. Introduction

The idea that strategic thinking—like other human cognitive capabilities—is not only importantly shaped by experience, but also that different kinds of experiences can offer different opportunities of strategic learning is certainly intuitive. Nonetheless, we currently have limited understanding of the basic mechanisms through which experience with previous interactive decisions shapes decision makers' strategic sophistication. In this study, we address this issue in the specific context of two-person, strategic-form games, in which strategic sophistication is usually intended as the ability of players to accurately predict the moves of their opponent based on the sole analysis of the payoff structure. Our research question is about *how* and under *what conditions* experience with previous decisions in two-person, 3×3 strategic-form games affects the degree of strategic sophistication of an individual.

To this end, we designed and ran an eye-tracking experiment consisting of three stages, in which players play sequences of games (with a unique pure-strategy Nash equilibrium) against an artificial counterpart that always selects the equilibrium strategy. In the initial stage, players' degree of strategic sophistication is assessed based on their lookup patterns in a sequence of games of different types (i.e., differing for their dominance structure) played without feedback. In the second stage (called Learning stage), players play, with complete feedback after each choice, another sequence of games of one of the types encountered in the initial stage (the game type being our between-subject factor varying across treatments). This second stage was designed to test the effects of experience with games with different dominance structures on players' strategic sophistication. Finally, in the third stage, players' degree of strategic sophistication is reassessed analogously as in the first stage.

Our contribution is four-fold. We show that:

1. Strategic sophistication can be significantly enhanced by the experience gained in the Learning stage.
2. The newly acquired strategic capabilities are generalized to the types of games that were not faced in the Learning stage.
3. The type of games included in the Learning stage crucially affects players' ability to select the equilibrium strategy, and can improve or hinder strategic learning independently from players' initial degree of strategic sophistication. The game types included in the treatments offer substantially different learning opportunities to players, even though the behavior of their counterpart is invariantly and unambiguously that of selecting the equilibrium strategy.
4. Different types of players learn in substantially different ways. Among those players who learn to play the equilibrium strategy in the games faced during the Learning stage, only one group of players display a strategic intent compatible with best responding to the opponent's incentives, and successfully transfer such learning to other types of games. Instead, the other players appear to learn some heuristics of play which are also generalized to other game types, but that lead less often to the equilibrium strategy.

With the third and fourth points of our contribution, we highlight boundary conditions for the effects of learning on strategic sophistication. Indeed, it is reasonable to assume that different game situations offer players different learning opportunities: Whereas some game structures can induce players to focus on their own payoffs, others can induce them to reason more strategically about the payoffs of their opponent, even when feedback about their opponent's choices is provided in both situations. For example, games in which players have a dominant strategy belong to the former category, because choice behavior is independent from any beliefs about the opponent's choices (see, for example, Brandenburger, 1992). Consequently, repeated experience with these games is expected to reinforce the tendency to focus on one's own payoffs and to neglect those of the other. Instead, games in which players do not have a dominant strategy are expected to induce them to also analyze their opponents' incentives, in the attempt to anticipate the other player's moves.

In our study, we use the eye-tracking methodology, as the analysis of the lookup patterns that precede the selection of an action allows for a reconstruction, at the individual level, of the payoff information gathering process.[1] The lookup patterns we consider in our analysis have been demonstrated to be strongly connected with the underlying decisional process (Arieli et al., 2011; Glöckner and Herbold, 2011; Polonio et al., 2015; Graffeo et al., 2015; Devetag et al., 2016; Chen et al., 2018; Polonio and Coricelli, 2019; Zonca et al., 2019, 2020a). Therefore, the analysis of eye-tracking data allows us to untangle changes in the strategic sophistication of players from changes in their beliefs about the actions of their opponent, or from changes in their beliefs over the type of opponent they believe they are matched with. This would be hard to achieve solely based on the analysis of choice data. In this way, we can observe whether experience with previous interactions translates into changes in the way players search for payoff information—a reliable proxy of how they *elaborate* such information.

So far, the eye tracking technique[2] has been mainly used in behavioral game theory research for assessing individuals' strategic sophistication in "static settings," i.e., in decision problems intentionally designed to limit as much as possible any learning effects (i.e., one-shot, two-person games played without feedback, as in Devetag et al., 2016). Two exceptions to this line of research are the eye-tracking studies by Knoepfle et al. (2009), in which participants play sequences of two-person 4×4 games with feedback, and that by Zonca et al. (2019), in which participants are taught to become more strategic in two-person 2×2 games by being shown alternative, more sophisticated rules of play. However, these two studies leave unanswered the question of how strategic learning evolves with experience, and how individuals spontaneously extract and make use of the knowledge gained in previous experiences with interactive decisions (for a theoretical account and an experimental test of this generalization tendency, see Grimm and Mengel, 2012, and Mengel, 2012). Relatedly, economics and psychology studies of learning in games have mainly focused on the attempt to replicate the (aggregate) process of mutual adaptation of players' choice behavior (e.g., Erev and Roth, 1998; Camerer and Ho, 1999; Marchiori and Warglien, 2008), but without investigating how experience shapes strategic reasoning.

Heterogeneity of strategic sophistication across individuals is a well-established fact: Previous studies have shown that sophistication is not only linked to and conditioned by a variety of individual factors, such as cognitive skills and psychological traits (e.g., Devetag and Warglien, 2003; Carpenter et al., 2013; Bayer and Renou, 2016; Gill and Prowse, 2016; Proto et al., 2019), but also affected by the type and representation of the decision problem (Georganas et al., 2015). In addition, evidence exists that individuals can endogenously adjust their own level of sophistication based on expectations about that of their counterparts (Agranov et al.,

[1] A further advantage of the eye-tracking technique is its ecological validity. Experiments run with and without recording eye movements have highlighted no differences in choice behavior (Wang et al., 2010; Polonio et al., 2015; Polonio and Coricelli, 2019).

[2] The same applies to the other elicitation techniques. Several other methods for eliciting strategic sophistication have been proposed in the literature, each associated with advantages as well as disadvantages (see overview and discussions in Franco-Watkins and Johnson, 2011, and Mauersberger and Nagel, 2018).

2012; Slonim, 2005), or on cost/benefit analyses (Alaoui and Penta, 2015). The study by Gill and Prowse (2016) importantly shows how cognitive skills and personal traits affect how people learn to play equilibrium in p-Beauty contest games, with high-cognitive ability agents being able to reach equilibrium faster and earning substantially more than low-cognitive ability ones. Our contribution goes further along this direction, by highlighting another determinant factor for heterogeneity of strategic sophistication—i.e., the nature of strategic experiences people are exposed to—, and showing how this factor interacts with people's default strategic approach.

2. Experimental design and procedure, and definition of gaze transitions

2.1. Experimental design

To answer the research questions presented in the introduction, we designed an experiment in three stages (see Fig. 1 for an overview of the experimental design). In each stage, participants played, as row player,[3] a sequence of two-person 3×3 games in strategic form with a unique equilibrium in pure strategies.

Participants were matched against an algorithm (henceforth, the *Computer*) that always played the equilibrium strategy. At the beginning of the experiment, participants were informed that the Computer would play rationally, trying to maximize its own payoff, and that it would not modify its strategy during the experiment, nor adjust its choices to those of its counterpart (see instructions in section S.9 of the Supplementary material). Such strategy is clearly identifiable in the Learning stage of both feedback treatments of our experiment (as discussed later). The choice of matching participants against an artificial opponent aligns subjects' beliefs and eliminates the problem of uncontrolled endogenous adjustments of the level of strategic sophistication that could arise when playing against a human counterpart (see Agranov et al., 2012). Such an enhanced control comes at the cost of limiting the generalizability of our results: Interacting with an artificial counterpart could trigger strategic considerations and reasonings different than those arising in interactions between humans.

We now describe in detail the three stages of our experiment.

The Assessment Stage: Stage 1, which we call *Assessment*, was designed to assess the initial level of strategic sophistication of each participant. In the Assessment, participants play without feedback a sequence of 15 two-person 3×3 games, G_t^A $(t = 1, \dots, 15)$: Five games have a dominant strategy for the row player, i.e., the participant ("Dominant-Self" games, henceforth DS games), but no dominant strategies for the column player; five games have a dominant strategy for the column player, i.e., the Computer ("Dominant-Other" games, henceforth DO games), but no dominant strategies for the row player; and, finally, five games have no dominant strategies for either of the players, and are not solvable through iterated dominance ("No-Dominance" games, henceforth ND games).

The Learning Stage: In stage 2, which we call the *Learning* stage, players play 20 different instances of either a DS or DO game, with and without feedback information after each choice in a 2 (DS/DO) × 2 (feedback/no-feedback) between-subject design. The games in this stage, G_t^L $(t = 1, \dots, 20)$, have been obtained by adding independent random (positive and negative) constants to the payoffs of one original game G_0^L (either a DO or a DS game, depending on the

[3] Previous research adopting a similar experimental design shows no differences in behavior between playing as row or column player (Polonio et al., 2015).

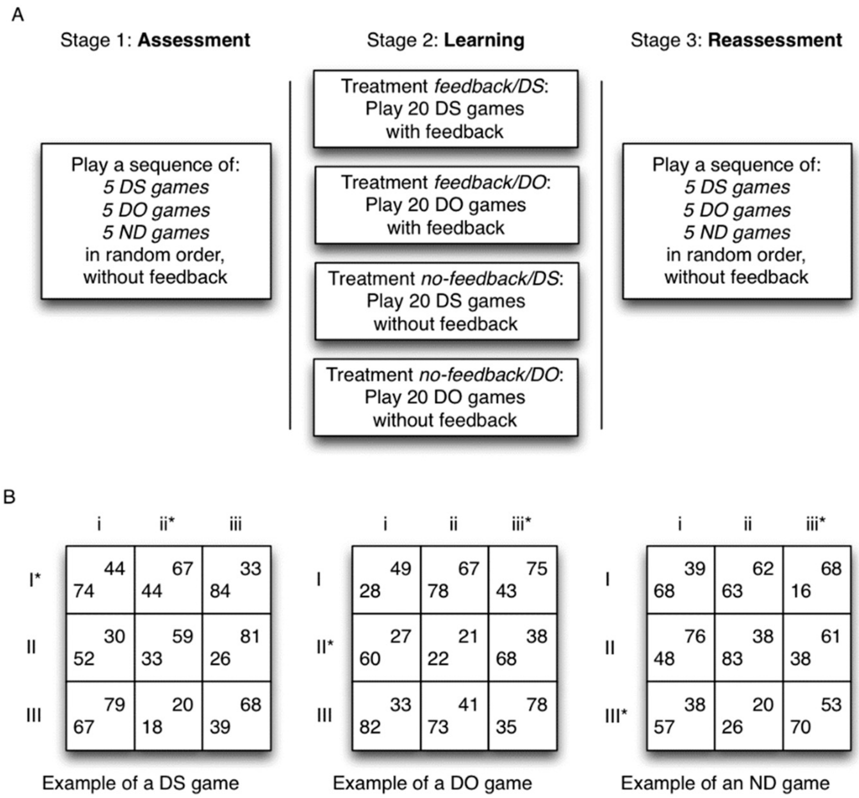

Fig. 1. Graphical summary of the experimental design (A), and instances of "dominant-self" (DS), "dominant-other" (DO), and "no-dominance" (ND) games (B). In Panel B, asterisks "*" indicate equilibrium strategies. In our three-stage experiment, the Assessment and Reassessment stages are identical across treatments. Treatments differ for the type of games (2 levels: DO/DS) and feedback (2 levels: with/without) in the Learning stage.

treatment), maintaining unaltered the equilibrium and dominance features of the original game. In addition, the order of rows and columns of each G_t^L game was randomly and independently permuted. Finally, also for G_t^L games, payoffs were in the interval $[10, 99]$.

The Reassessment Stage: In stage 3, which we call *Reassessment*, players played without feedback another sequence of 15 two-person 3×3 games, G_t^R $(t = 1, \ldots, 15)$. The payoffs of each G_t^R game were obtained by adding independent random (positive and negative) constants to the payoffs of the corresponding G_t^A game, without altering the game's equilibrium and dominance structure. For both G_t^A and G_t^R games, the order of rows and columns was randomly and independently permuted, and the order of games randomized. In all games, payoffs were within the interval $[10, 99]$. We designed stage 3 to observe how experience gained in stage 2 affects the individual level of strategic sophistication. All games used in the three stages are reported in Section S.10 of the Supplementary material.

Strategic salience of DS, DO, and ND games: From a behavioral perspective, DS, DO, and ND games can be associated to an increasing degree of strategic salience. That is, repeated play of these games differently reinforces the tendency to consider the incentives of the opponent. Because of the presence of a dominant strategy for the row player (the role played by participants), recognition of the equilibrium strategy in DS games requires a comparatively lower analytical effort: In these games, equilibrium play does not require awareness of the incentives of the opponent, nor forming beliefs about the actions of the counterpart (Brandenburger, 1992; Aumann and Brandenburger, 1995).[4] In addition, in DS games, because of the absence of a dominant strategy for the Computer, trying to predict its moves is more difficult. This could also discourage some players (especially those less sophisticated) from inspecting systematically the payoffs of their opponent and encourage them on focusing on their owns.

Instead, in DO games, equilibrium play does require the analysis of the counterpart's payoffs. In DO games, only the players that examine the full structure of incentives are expected to recognize their opponent's dominant action and best reply to it. It is worth noting that DO games were designed in such a way that the action with the largest payoff sum for the row player (which would be attractive for players that mostly focus on their own payoffs) *never* corresponds to the equilibrium strategy. In addition, in DO games, the absence of a dominant strategy for the row player makes it difficult to individuate an obvious choice. This could also encourage players to systematically inspect the payoffs of the Computer in the attempt to best respond to it. In this sense, in these games the strategic dimension is more salient than the DS ones.

Despite the difference in strategic salience, repeated play with complete feedback of both DS and DO games in the Learning stage should in principle allow participants to become more strategic, as in both situations players can unambiguously observe the best-responding strategy of the Computer. However, in DS games, the presence of an own dominant strategy and the absence of a dominant strategy for the opponent is expected to reinforce the participants' tendency (the default one for some players) to focus mostly on their own payoffs.

Equilibrium play in ND games requires yet a deeper analysis of the incentive structure. In ND games there are no dominant nor strictly dominated actions for either of the two players, so that equilibrium cannot be found via the method of iterated dominance: This makes the recognition of the equilibrium strategy cognitively more difficult in these games than in the DO ones. We chose to include ND games in the Assessment and Reassessment primarily to test the extent to which the strategic skills acquired in the training with DS and DO games (see the description of the Learning stage) are transferred to similar but not identical games. This point is further addressed in the Results section.

We label as *treatment DS* the treatments in which the Learning stage includes DS games, and as *treatment DO* those in which the Learning stage includes DO games. *No-feedback* and *feedback* treatments are introduced to control for the effect on learning of immediate and complete feedback information from that of mere repeated game play.

Combined, the four treatments defined by the Learning Stage allow us to assess the effect of learning on strategic sophistication, as well as to draw boundary conditions for this phenomenon. It is worth noting that higher levels of strategic sophistication are generally associated with a higher analytical effort (other than a higher cognitive difficulty), which in general is increasingly costly to exert. In order to limit as much as possible the effect of cost/benefit considerations on the exerted level of sophistication (see Alaoui and Penta, 2015), all games are constructed in such a way that the average payoff differs considerably between the participants that behave consistently with best-responding to uniform beliefs over $a_{-i},\ i = 1, 2, 3$ (45.3 points)[5] and those that select the equilibrium strategy (65.6 points; the difference is the same in all treatments).

[4] The ease of finding the dominant strategy can be enhanced by representing the game in extensive form as shown by Brocas et al., 2018.

[5] Or that simply select the action with the largest payoff sum. In both cases, these players do not invest much cognitive effort in analyzing the incentives of their opponent.

2.2. Experimental procedure

Two hundred forty-three subjects took part in the experiment (81 males, $M_{\text{age}} = 24.1$, $SD_{\text{age}} = 4.6$): 91 in treatment *feedback/DS*, 89 in treatment *feedback/DO*, 32 in treatment *no-feedback/DS*, and 31 in treatment *no-feedback/DO*. We originally aimed at collecting data for 30 participants per treatment, in alignment with similar experimental studies documented in the literature (e.g., Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006; Brocas et al., 2014; Polonio et al., 2015; Devetag et al., 2016; Polonio and Coricelli, 2019). However, in order to adequately increase the size of the groups obtained with the cluster analysis, we ran an additional data collection for the feedback treatments. For both data collections, data were analyzed only upon completion of the data collection.

The experiment was run with one participant at a time. Upon arrival, each participant was given a hard copy of instructions, which were then read aloud by the experimenter. Each participant had to complete a questionnaire designed to test the comprehension of the task. If the questionnaire was incorrectly answered or under the request of the participant, instructions were repeated until full understanding. No participants failed the comprehension check more than twice. At each trial, participants selected their preferred action by pressing either the “1,” “2,” or “3” key of the keyboard. Payoffs in each matrix were represented in different colors for different players (see Fig. 2A) to enhance comprehension (there were no color-blind participants). In the feedback treatments, participants could inspect the feedback screenshot (appearing after each choice) displaying the whole payoff matrix with arrows indicating their own and the Computer’s choice, with the corresponding matrix cell highlighted by a green border (see experimental instructions in Section S.9 of the Supplementary material). Participant could inspect the feedback screenshot without time constraints. In all experimental stages, the order of the games was independently randomized for each participant.

At the end of the experiment, one trial per stage was randomly selected, and the participant was paid based on performance in the selected trials. The experiment, conducted at the Experimental Psychology Laboratory of the University of Trento (Italy), lasted about one hour, and the average payoff was about 16 euros (about \$16.4).

Participation was on a voluntary basis and participants who had participated in previous eye-tracking experiments on interactive decisions were not allowed to participate. We did not make any ex-post data exclusions and present in the paper all data we collected.

2.3. Codification and interpretation of eye tracking data

To analyze eye movements, we defined eighteen Areas of Interest (AOIs) centered on the matrix payoffs (see Fig. 2B). AOIs have a circular shape with an area of 36,000 pixels. AOIs cover 46.8% of the game matrix area and are not overlapping. Although the AOIs only cover a small portion of the entire visual space, almost all fixations (92%) are located within the AOIs. Fixations not falling within any of the AOIs are discarded from the analysis.

For the eye-tracking data analysis, we mostly focus on types of *transitions*, rather than on mere *fixations*. Fixations[6] allow to infer the share of attention allocated to one’s own and the other’s

[6] Eye fixations occur when a participant keeps his/her visual gaze still. In this study, a fixation occurs when the gaze is focused within 1° of visual angle for at least 80 ms, a threshold commonly adopted to discriminate between fixations and other types of ocular activities.

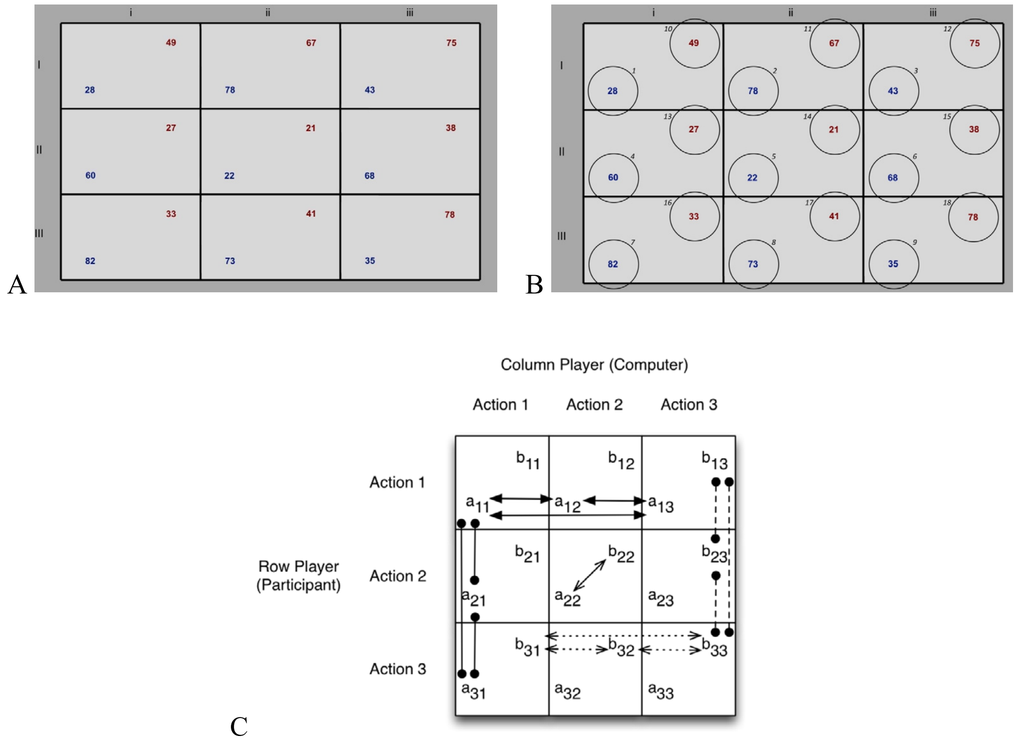

Fig. 2. Panels A and B illustrate, respectively, examples of the stimuli presented to players and of the definition of the 18 areas of interest (AOIs). Panel C illustrates examples of the five types of transitions defined in our analysis: solid line arrows illustrate examples of *own-sum* transitions; solid line circles exemplify *own-dom* transitions; dashed line circles *other-sum* transitions; dotted line arrows *other-dom* transitions; and solid line empty arrows *intracell* transitions. The direction of transitions from one AOI to another is irrelevant for classification. (For interpretation of the colors in the figures, the reader is referred to the web version of this article.)

payoffs. Instead, transitions (defined as consecutive fixations that fall within two different AOIs) provide information about how participants integrate payoff information, a process that cannot be reconstructed by the analysis of fixations alone. According to the classification of transitions proposed by Devetag et al. (2016), and Polonio and Coricelli (2019), we consider the following five types of transitions for the *row player* (cf. the game matrix illustrated in Fig. 2C).

1) Own-payoff within-strategy transitions (henceforth *own-sum*), which connect the AOIs centered on payoffs a_{ij} and a_{ik}, with $i, j, k = 1, 2, 3$. These transitions are typically used by people to detect their own strategy with the largest payoff sum (or average).

2) Own-payoff between-strategy transitions (henceforth *own-dom*), which connect the AOIs centered on payoffs a_{ij} and a_{kj}, with $i, j, k = 1, 2, 3$. These transitions are typically used by people to detect their own dominant strategy, or find a best response to the predicted action of the counterpart (see the temporal analysis of transitions in section S.1 of the Supplementary material).

3) Other-payoff within-strategy transitions (henceforth *other-sum*), which connect the AOIs centered on payoffs b_{ij} and b_{kj}, with $i, j, k = 1, 2, 3$. These transitions are typically used by people to identify their opponent's strategy with the largest payoff sum (or average).

4) Other-payoff between-strategy transitions (henceforth *other-dom*), which connect the AOIs centered on payoffs b_{ij} and b_{ik}, with $i, j, k = 1, 2, 3$. These transitions are typically used by people to identify any dominant strategies of their opponent.

5) Intracell transitions (henceforth *intracell*), which connect the AOIs centered on payoffs a_{ij} and b_{ij}, with $i, j = 1, 2, 3$. These transitions are typically used by people to identify a focal point (as, for example, the cell that yields the largest payoff sum), or similar strategies based on intracell payoff comparisons.

The association of transitions to the different strategic intents has been demonstrated in experimental settings involving 2×2 two-person games in the strategic form. These experiments show that when participants are instructed to choose based on a given decision rule (e.g., finding out the dominant actions in a game, or the action associated with the largest payoff sum), they actually employ the search-specific gaze transitions described above (see Polonio et al., 2015, and Zonca et al., 2019). Furthermore, the link between the type of transitions used and choice behavior has been extensively demonstrated in both 2×2 and 3×3 game matrices (e.g., Devetag et al., 2016; Polonio and Coricelli, 2019; and Zonca et al., 2020b).

3. Aggregate-level results

In this section, we first present and discuss aggregate choice data, and then aggregate eye-tracking data. In the next section, we propose an individual-level analysis of both choice and eye-tracking data, and which relies on a clustering of participants based on their lookup patterns in the Assessment stage.

3.1. Analysis of aggregated behavioral choice data

To observe how experience affects choice behavior, we compare the frequency of equilibrium choices in the three game types (DS, DO, and ND) in the Assessment and Reassessment stage, in all four treatments (Fig. 3).

Looking at the two treatments with feedback (Fig. 3, top row), a two-way repeated-measure ANOVA (within-subject factor: *stage*; between-subject factor: *treatment*) indicates that immediate feedback significantly improves the frequency of equilibrium choices in all game types (cf. the significant effect of the *stage* factor, Table 1). The driver of this improvement in DO and ND games is the experience gained with DO games during the Learning stage (cf. the significant *treatment*stage* interaction, Table 1): Equilibrium choices in DO games pass from 46%, SD = .37, to 91%, SD = .22, whereas in ND games from 44%, SD = .35, to 70%, SD = .39 (the difference is significant in both cases: $t(88) = 11.344$, $p < .001$, for DO games, and $t(88) = 5.832$, $p < .001$, for ND games). This result suggests that experience gained with DO games during the Learning stage is also transferred to deal with the seemingly similar, but cognitively and strategically more demanding, ND games. As for the DS games, the improvement is instead due to experience gained with DS games during the Learning stage (cf. the significant *treatment*stage* interaction, Table 1): Equilibrium choices in DS games pass from 74%, SD = .33, to 88%, SD = .25, and the difference is significant ($t(90) = 3.612$, $p < .001$). The fact that experience with DO games enhances performance in ND games, but not DS games, could be explained by a simple ceiling effect: Since the equilibrium strategy is relatively easy to recognize in DS games (as they do not require any strategic reasoning), the proportion of equilibrium choices is already large in the Assessment, leaving little room for improvement. However, the cluster-level

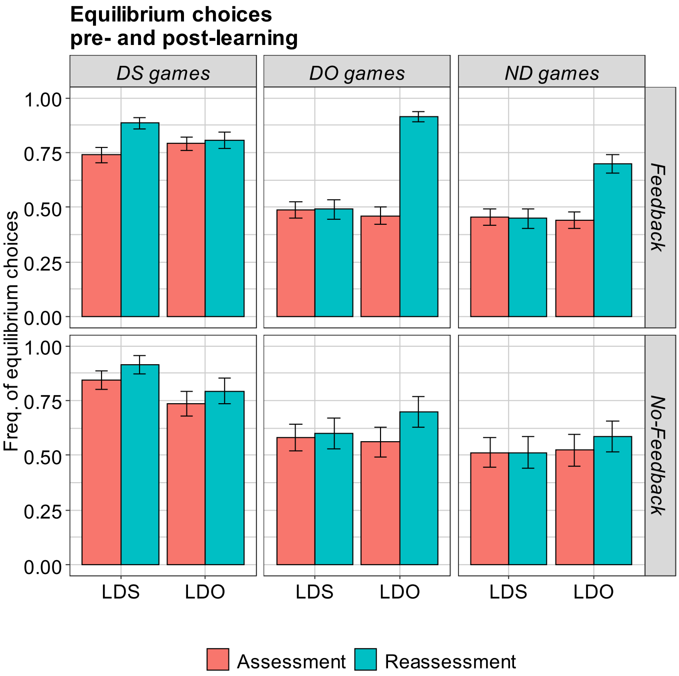

Fig. 3. Frequency of equilibrium choices in the Assessment and Reassessment, by treatment and game type. The label "LDS" indicates Learning with DS games, and "LDO" Learning with DO games. Error bars are ± 1SE.

analysis discussed later reveals that this is instead due to players' heterogeneous reactions to experience.

The treatments without feedback in the Learning stage show instead almost no learning effects (Fig. 3, bottom row). In both treatments, the absence of feedback in the Learning stage does not produce significant differences in the frequency of equilibrium choices across the Assessment and Reassessment stages in DO and ND games (cf. the non-significant effect of the *stage* factor, Table 1). Furthermore, no learning effect based on a specific treatment is observed (cf. the non-significant *treatment*stage* interaction, Table 1). Although the main effect of the *stage* factor is significant for DS games, indicating an improvement of equilibrium play, such improvement is not due to the kind of games faced in the Learning stage (cf. the non-significant *treatment*stage* interaction, Table 1). This means that repeated play of DO and DS games without feedback equally enhances equilibrium play in DS games.

In summary, behavioral results suggest that, within our framework, feedback is a necessary, but not sufficient, condition for the improvement of equilibrium play. It is instead the combination of feedback and type of experience (specifically, the one that emphasizes the role played by the other player) that produces a large increase of equilibrium play.

However, choice data do not show the cause of such an increase in the frequency of equilibrium choices. In order to better understand the nature of the strategic learning participants engage in and how this is modified by experience, we now turn to the analysis of eye-tracking data, which, through the analysis of the different kinds of lookup patterns (transitions), offers a closer view on participants' strategic intent.

Table 1
Tests for differences in equilibrium proportions across experimental stages.

Game type	Main effect *stage* (*Assessment* and *Reassessment*)	Main effect *treatment* (*own*- and *other-focus*)	Interaction effect *treatment * stage*
		Feedback treatments	
DS	$F(1, 178) = 9.367$ $p = 0.003$	$F(1, 178) = 0.103$ $p = 0.749$	$F(1, 178) = 5.952$ $p = 0.016$
DO	$F(1, 178) = 46.92$ $p < 0.001$	$F(1, 114) = 22.59$ $p < 0.001$	$F(1, 178) = 47.04$ $p < 0.001$
ND	$F(1, 178) = 13.46$ $p < 0.001$	$F(1, 178) = 6.376$ $p < 0.012$	$F(1, 178) = 15.26$ $p < 0.001$
		No-feedback treatments	
DS	$F(1, 61) = 4.322$ $p = 0.042$	$F(1, 61) = 3.064$ $p = 0.085$	$F(1, 61) = 0.031$ $p = 0.861$
DO	$F(1, 61) = 1.967$ $p = 0.166$	$F(1, 61) = 0.233$ $p = 0.631$	$F(1, 61) = 1.154$ $p = 0.287$
ND	$F(1, 61) = 0.447$ $p = 0.506$	$F(1, 61) = 0.226$ $p = 0.636$	$F(1, 61) = 0.461$ $p = 0.500$

Note. Two-way repeated measures ANOVAs of the frequency of equilibrium choices in the Assessment and Reassessment, for each of the three types of games (within-subject factor: *stage*; DS indicates dominant-self games; DO indicates dominant-other games; ND indicates no-dominance games), in the *DS* and *DO* treatments (between-subject factor: *treatment*).

3.2. Analysis of aggregated eye-tracking data

To evaluate the effect of experience on strategic sophistication, we analyze whether fixations and lookup patterns significantly mutate across the Assessment and Reassessment stages within a given treatment, and whether fixations and lookup patterns in the Reassessment differ across treatments. Fixations give a first hint on the payoff information gathering process, showing how attention is allocated to one's own and to the opponent's payoffs, and how such allocation of attention is affected by experience (Fig. 4).

As shown by Fig. 4, the distribution of attention across own and other payoffs is similar in the Assessment stage of the various treatments (and aligned with the literature, see Polonio and Coricelli, 2019). Only experience in the *feedback/DO* treatment results in a significant shift from own- to other-payoff fixations (from 47%, $SD = .22$, to 67%, $SD = .16$, $t(88) = 8.617$, $p < .001$). This result, together with the corresponding increase in equilibrium choices seen earlier, suggests an increase in the level of sophistication of the players exposed to repeated play of DO games with feedback. In all other treatments the shift from own- to other-payoff fixations is not significant.

However, fixation data only show how much attention is devoted to own and other's payoffs, but not how such attention is *used* by players. Transition data (or lookup patterns) can instead shed light on the qualitative dimension of attention, that is, how the information on own and the other's payoffs is collected—and then processed. In addition, looking at how the transition patterns change across stages, can inform us about how experience affects the way players analyze the game. Fig. 5 gives a graphical representation of the distribution of transitions by treatment and stage.

To understand how experience affects the way players analyze the game, for each participant and within the same experimental stage, we computed the proportion of the five types of

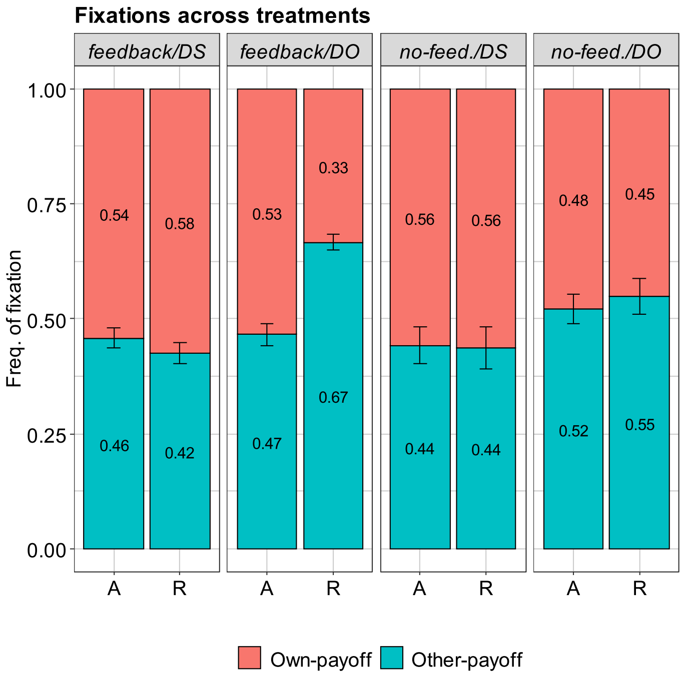

Fig. 4. Frequency of fixations in the Assessment and Reassessment, by treatment. The label "A" indicates the Assessment, and "R" the Reassessment. The label "Own-payoff" ["Other-payoff"] indicates the share of fixations on row player's [column player] payoffs. Error bars are ± 1SE.

transitions (cf. section 2.3), normalized by their sum (as done in Polonio et al., 2015).[7] Such normalization weighs equally individual observations and allows us to focus on those transition types that have a clear interpretation in terms of the underlying strategic analysis carried out by players. We refer to the ensuing vector of five components computed for each participant as a *transition distribution* (also commonly referred to as *compositions* in compositional analysis; see Aitchison, 1986). Table 2 reports the centers of transitions (i.e., the geometric means over all participants of each type of transition) in the Assessment and Reassessment stages of the four treatments, as well as their statistical comparisons across stages. The description of the statistical framework for the analysis of change of transition distributions is reported in the Supplementary material (see Aitchison, 1986; Pawlowsky-Glahn et al., 2015).

In the two *no-feedback* treatments (lower panel of Fig. 5), transition distributions do not change significantly across stages (see Table 2, column on the right). Thus, the experience gained in the Learning stage does not significantly alter the way in which participants analyze the payoff matrix.

Instead, in the two *feedback* treatments, experience does have a significant effect (see Table 2 and top panel of Fig. 5). In the *feedback/DS* treatment, passing from the Assessment to the Reassessment stage, *own-sum* transitions are decreased in favor of *own-dom* transitions, and in general less attention is paid to the opponent's payoffs (both *other-sum* and *-dom* transitions are decreased). Thus, participants appear to learn to detect dominance in their own actions.

7 The five types of strategic transitions do not include all possible types of eye transitions.

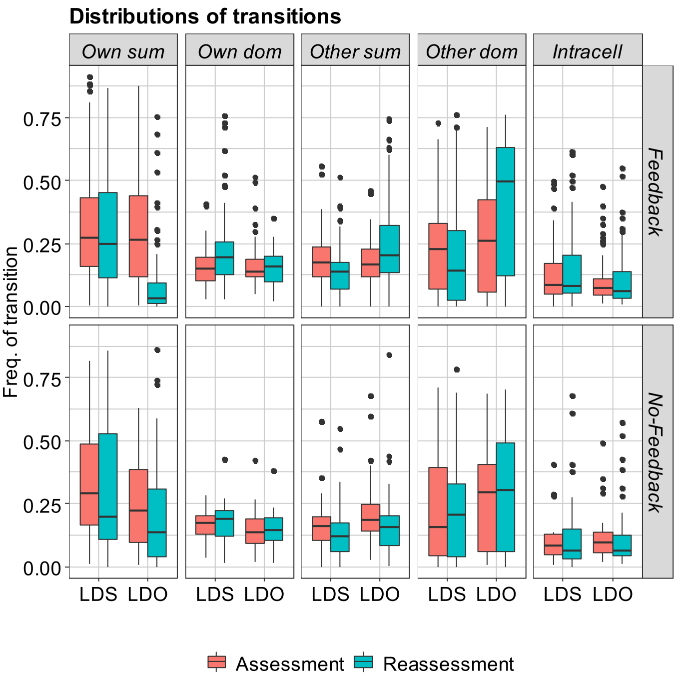

Fig. 5. Distributions of transitions by treatment and stage. The five columns represent the five types of transitions. The first row shows the distribution of transitions for the two treatments with feedback, and the second row the distribution of transitions for the two treatments without feedback. “LDS” and “LDO” indicate the Learning stage in the treatments including, respectively, DS and DO games. Boxplots visualize the median, two hinges (spanning the inter-quartile range, or the distance between the first and third quartiles, indicated with IRQ), the whiskers, and all outlying points. The upper whisker extends from the corresponding hinge to the largest value no further than 1.5 * IQR from the hinge. The lower whisker extends from the corresponding hinge to the smallest value at most 1.5 * IQR of the hinge.

Table 2
Centers (i.e., geometric means) of transitions in the Assessment and Reassessment, by treatment, and statistical test of the difference.

Treatment	Stage	Centers of transitions					Signif. (Hotelling’s T2)
		own-sum	own-dom	other-sum	other-dom	intracell	
feedback/DS	A	.336	.211	.162	.161	.130	T2 = 4.867
	R	.293	.296	.142	.122	.147	Df = (4, 87)
							p = .001
feedback/DO	A	.313	.230	.167	.174	.117	T2 = 22.894
	R	.045	.203	.278	.372	.100	Df = (4, 85)
							p < .001
no-feedback/DS	A	.323	.226	.178	.158	.114	T2 = 2.226
	R	.309	.303	.110	.152	.125	Df = (4, 28)
							p = .092
no-feedback/DO	A	.223	.169	.254	.223	.130	T2 = 1.434
	R	.188	.213	.206	.260	.131	Df = (4, 27)
							p = .250

Note. In the column named “Stage”, “A” indicates the Assessment, and “R” the Reassessment. Due to rounding, the reported distributions may not sum up exactly to one.

In the *feedback/DO* treatment, experience has a different impact on players' strategic sophistication. In this treatment, experience with the strategically more salient DO games induces participants to allocate more attention to the inspection of their counterpart's incentives, at the expense of *own-sum* transitions, which are reduced from 31% (SD = .23) in the Assessment to 9% (SD = .15) in the Reassessment. Instead, *other-dom* transitions, used to detect dominance in the counterpart's actions, increase from 26% (SD = .21) to 41% (SD = .25) as well as *other-sum* transitions (from 17%, SD = .10, to 25%, SD = .17), used to detect the opponent's strategy with the largest payoff sum.

As for the comparisons across treatments, we are interested in testing whether the experience in the *feedback/DO* treatment alters transition distributions in a significantly different way than the other treatments. More specifically, we ask ourselves whether the distributional changes (referred to as *perturbations* in compositional analysis) from the Assessment to the Reassessment are significantly different across treatments. The maximum likelihood procedure used for comparing independent samples of perturbations (described in the Supplementary material) answers this question positively. Comparing the treatments *feedback/DS* and *feedback/DO*, transition perturbations differ in both the center and variance structure (we reject H_{03} in favor of H_{gen}; $X^2(4) = 34.42$, $p < .001$; see Supplementary material for the description of the hypothesis structure). As for the comparisons of *feedback/DO* with the no-feedback treatments, transition perturbations differ in the centers, although not in the variance structure (we reject H_{01}, but fail to reject H_{02}; $X^2(14) = 17.42$, $p = .23$, for the comparison with *no-feedback/DO*; $X^2(4) = 12.17$, $p = .59$, for the comparison with *no-feedback/DS*). Thus, the significant differences in the perturbation centers show that the effect of experience on transition distributions in the *feedback/DO* treatment is significantly different from that observed in all other treatments.

Together, these results from the analysis of aggregate eye-tracking data corroborate the interpretation of choice data presented earlier. Only the combination of immediate feedback and experience in a strategic decision setting that emphasizes strategic interdependence of choices results in a more complete analysis of the payoff matrix, which, in turn, enhances equilibrium play. It is important to note that the relation between lookup patterns and frequency of equilibrium play is a relation of causality, the former phenomenon causing the latter, as demonstrated in Polonio et al. (2015). Participants not only allocate relatively more attention to the other's incentives, but also make use of those gaze transitions (especially *other-dom* transitions) that are consistent with equilibrium reasoning. This important point will be further elaborated in the individual analysis sections, and it is also supported by the temporal analysis of gaze transitions reported in the Supplementary material. In addition, in the *feedback/DO* treatment, participants developed some generalizable experience that they also applied to optimally play games on which they were not trained, i.e., the non-dominance solvable ND games.

Overall, these results provide evidence that the level of strategic sophistication of an individual is plastic, and that it can be substantially enhanced by experience. The individual analysis that follows gives a more insightful picture of players' heterogeneous learning styles.

4. Individual-level results

In this section, we propose an analysis of individual-level data that relies on the clustering of participants. Such clustering is based on the lookup patterns (transitions) observed in the Assessment stage of the experiment. We first discuss how clusters are constructed, then present and discuss cluster-level choice data, and, finally, cluster-level eye-tracking data.

Table 3
Number of participants in each cluster by treatment.

	feedback/DS	feedback/DO	no-feedback/DS	no-feedback/DO	Total
Cluster 1	47	54	16	17	**134**
Cluster 2	19	22	10	6	**57**
Cluster 3	25	13	6	8	**52**
Total	**91**	**89**	**32**	**31**	**243**

4.1. *Definition of the clusters*

We analyze how different player types react to experience. Specifically, our analysis shows which groups of participants are the drivers of the aggregate effect of experience on transitions (illustrated in Fig. 5), but also evaluates the possibility that different types of participants engage in fundamentally different types of learning. Indeed, the increase of frequency of equilibrium choices highlighted in the aggregate analysis could in principle mask underlying reasonings that depart from equilibrium reasoning (see, for example, Crawford et al., 2013).

To this end, we carry out a cluster analysis of participants based on the distribution of transitions recorded in the Assessment, pooling together the data for the *feedback* and the *no-feedback* treatments (this because the Assessment stage is common across all four treatments). We use the Gaussian Mixture Modeling approach for model-based clustering described by Fraley and Raftery (2002). A point of strength of this approach is that the clustering model that best describes data is endogenously determined by the procedure: Clustering models are first estimated via the *Expectation-Maximization* algorithm, and then selected based on the *Bayesian Information Criterion* (see details in section S.8 section of the Supplementary material).

Consistently with previous research (Polonio et al., 2015; Devetag et al., 2016), the best clustering model is the one that categorizes participants into three clusters as reported in Table 3.

4.2. *Analysis of clustered behavioral data*

We here analyze the frequency of equilibrium choices in the Assessment and Reassessment, by cluster and treatment (see Fig. 6). In this and the following sections, we focus on data from the feedback treatments, as the no-feedback treatments have been previously shown not to affect substantially strategic sophistication, irrespective of the games included in the Learning stage. In addition, the disaggregated analysis of choice behavior in the no-feedback treatments (see Supplementary material for the details) confirms the very limited effect of the Learning stage on strategic sophistication, thus not revealing effects that cancel out with the aggregation of data.

In the *feedback/DS* treatment (marked as "LDS" in Fig. 6), the frequency of equilibrium choices averaged over all three types of games does not change significantly across the Assessment and Reassessment in any of the clusters (paired t-tests for the equilibrium frequencies averaged over all types of games: $t(46) = 0.56$, $p = .58$, for the initially strategic players in Cluster 1; $t(18) = 1.10$, $p = .29$, for players in Cluster 2; and $t(24) = 1.61$, $p = .12$, for players in Cluster 3), reflecting the result seen for the aggregate data. However, Cluster 1 and 3 appear to significantly increase equilibrium play in DS games (from 83%, $SD = .24$, to 93%, $SD = .22$, $t(46) = 2.535$, $p = .015$, for Cluster 1; from 50%, $SD = .38$, to 86%, $SD = .27$, $t(24) = 3.662$, $p = .001$, for Cluster 3). We will see that this latter result is supported by changes in the distribution of transitions that will be discussed in the next session.

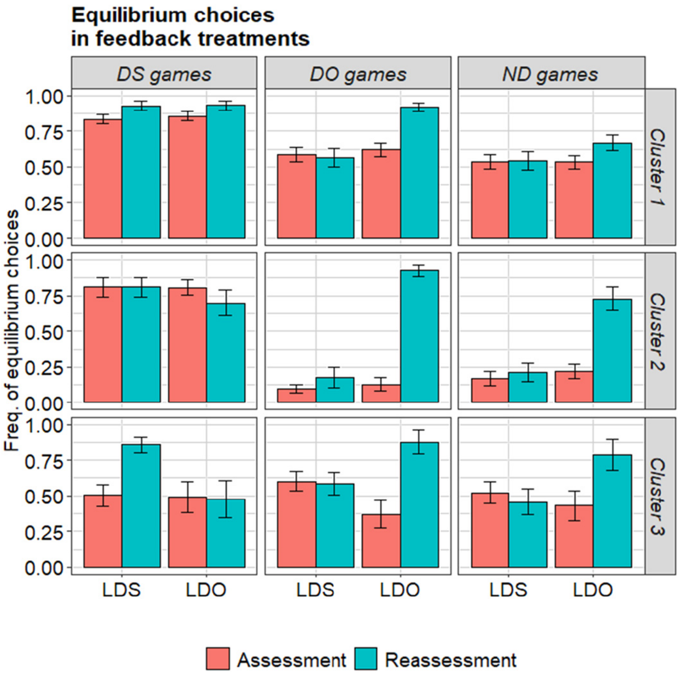

Fig. 6. Frequency of equilibrium choices by cluster, game, and stage in the feedback treatments. The label "LDS" indicates Learning with DS games, and "LDO" Learning with DO games. Error bars are ± 1SE.

Instead, experience in the *feedback/DO* treatment significantly increases the frequency of equilibrium choices for all clusters (paired t-tests for the equilibrium frequencies averaged over all types of games: from 67%, SD = .23, to 84%, SD = .17, t(53) = 6.79, $p < .001$, for the Cluster 1; From 38%, SD = .07, to 78%, SD = .20, t(21) = 8.58, $p < .001$, for Cluster 2; and from 43%, SD = .31, to 71%, SD = .28, t(12) = 3.18, p = .008, for Cluster 3). Even though starting from very different frequencies of equilibrium choice in the Assessment stage, players in the three clusters similarly learn to select the equilibrium choice in the DO and ND games in the final Reassessment stage (83%, SD = .24, for the Cluster 1; 79%, SD = .24, for Cluster 3; and 83%, SD = .32, for Cluster 3; the difference across clusters is not significant, F(2, 86) = 0.216, p = .81). Nonetheless, players in Cluster 2 and 3 fail to match the frequency of equilibrium choices by Cluster 1 players in the simple DS games, in which dominance in the row player's actions should instead facilitate the detection of the equilibrium choice. In the DS games of the Reassessment, players in Cluster 2 and 3 select, respectively, the equilibrium option only 70% (SD = .43) and 48% (SD = .47) of the time, whereas for Cluster 1 the frequency of equilibrium choices is 93% (SD = .22). The difference is significant (F(2, 86) = 11.75, $p < .001$), and only the pairwise comparison of equilibrium frequency in DS games between Cluster 2 and Cluster 3 is not significant (for Cluster 2 vs. Cluster 3, Tukey's adjusted-p = .13; for Cluster 1 vs Cluster 2, Tukey's adjusted-p = 0.017; and for Cluster 1 vs. Cluster 3, Tukey's adjusted-$p < .001$). This result prompts the question of whether the learning that Cluster 1 players engage in is the same of those in Cluster 2 and 3. To answer this question, we turn to the analysis of individual eye-tracking data.

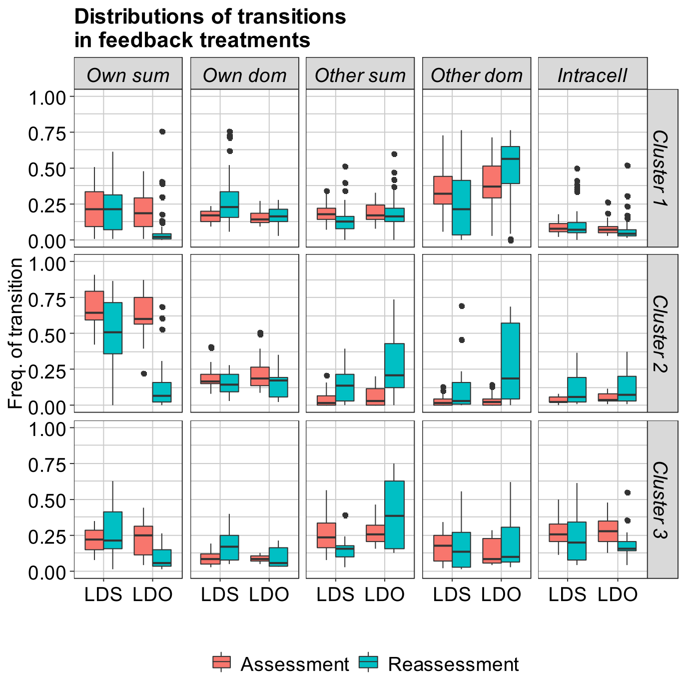

Fig. 7. Distributions of transitions in the *feedback* treatments, by cluster and stage. The five columns represent the five types of transitions, whereas the three rows the three clusters. "LDS" and "LDO" indicate the Learning stage in the treatments including, respectively, DS and DO games. Boxplots are constructed as in Fig. 5.

4.3. Analysis of clustered eye-tracking data in the Assessment stage

We begin discussing the distributions of transitions observed in the different clusters during the Assessment Stage, before any learning took place (see red bars in Fig. 7). Since clusters were created based on gaze transitions in the Assessment stage by pooling all participants together, we observe no differences in the distributions of transitions across the *feedback/DS* and *feedback/DO* treatments. Therefore, our observations about transitions in the Assessment stage refer to both treatments.

Cluster 1: Players carry out a comparatively more complete analysis of players' incentives, allocating a considerable share of attention to the inspection of their opponent's payoffs. Participants in this cluster make a larger use of *other-dom* transitions (37%, SD = .17), mainly at expense of *own-sum* transitions (20%, SD = .14). Such a lookup pattern identifies players that are since the beginning more strategically sophisticated (as in Costa-Gomes et al., 2001), and is compatible with best-responding to the opponent's incentives.

Cluster 2: Players mostly focus their attention on their own payoffs, displaying a larger frequency of *own-sum* transitions (63%, SD = .15), typically used to recognize the action with the largest payoff sum. Overall, participants in this cluster pay little attention to the payoffs of their opponent (the average of *other-sum* and *other-dom* transitions is 5%, SD = .06). Nonetheless, we cannot exclude that this cluster could also include some strategic players selecting the strategy that best responds to uniform beliefs over the Computer's actions $a_{-i},\ i = 1, 2, 3$ (analogous to a level-1 strategy; cf. Nagel, 1995), and that gather payoff information according to this belief (Fehr and Huck, 2016).

Cluster 3: Similarly to players in Cluster 2, players in this cluster appear to behave non-strategically at the beginning of the experiment, although they use yet another approach of analysis of the game payoffs. These participants frequently compare payoffs within a single matrix cell, i.e., through *intracell* transitions (27%, SD = .11).[8] These players do not analyze the game in a strategic way, mostly making *own*- and *other-sum* transitions (respectively, 22%, SD = .10, and 28%, SD = .13), beside the *intracell* ones.

Participants in Cluster 2 and 3 display an initial different analytical approach of the payoff matrix. However, a combined analysis of the behavioral and eye-tracking data in the Reassessment stage suggests a similar reaction to experience of players in these two clusters, and this reaction appears to be substantially different than that exhibited by players in Cluster 1.

4.4. Comparison of clustered eye-tracking data across Assessment and Reassessment

Comparing the transition distributions observed in the Assessment and Reassessment stages of the two feedback treatments, we can test whether all players respond to experience in the same way. The results of the statistical tests for these comparisons, by cluster and treatment, are reported in Table 4.

Lookup patterns of participants are significantly affected by experience in feedback treatments (cf. the column on the right of Table 4 and Fig. 7), meaning that transition distributions change significantly across the Assessment and Reassessment stages. In general, experience in the *feedback/DS* treatment increases players' tendency to focus on their own payoffs. In fact, players in Cluster 1 and 3 substantially increase the frequency of *own-dom* transitions at the expense of transitions focusing on *other* payoffs (the former by halving the frequency of the strategic *other-dom* transitions, the latter reducing the frequency of *other-sum* transitions). This change in the lookup pattern is consistent with the increase of equilibrium play in DS games that we highlighted in the previous section, when discussing cluster-level choice data. For players in Cluster 2, experience in this treatment lowers the frequency of *own-sum* transitions by slightly increasing that of *other-sum* ones. However, for these players, *own-sum* transitions still constitute about half of all transitions and this small increase in the attention allocated to their opponent's payoffs does not translate into a significant increase in the frequency of equilibrium choices.

Experience in the *feedback/DO* treatment has a different effect on transitions. In the Reassessment, Cluster 1 players remarkably increase the use of *other-dom* transitions that are useful to detect dominance in the opponent's actions, and decrease the use of *own-sum* transitions (they also keep using *own-dom* transitions to check for dominance in their actions and best reply to the expected strategy of their counterpart; see the temporal analysis of transitions discussed in section S.1 of the Supplementary material). This is a clear signal of an improved strategic sophistication, consistent with best-response reasoning. For players in Clusters 2 and 3, experience with DO games mainly translates into a decrease in *own-sum* transition, and an increase in *other-sum* and *other-dom* transitions, where *other-sum* transitions remain substantially more frequent than *other-dom* transitions. In addition, players in Cluster 2 still distinguish themselves for the

[8] Significantly more frequently than participants in the other two clusters. For players in Cluster 1 and 2, the average frequency of intracell transitions is 8% (SD = .05). The difference between this average frequency and that for players in Cluster 3 is statistically significant (Welch corrected $t(54.972) = 12.979$, $p < .001$).

Table 4
Centers (i.e., geometric means) of transitions in the Assessment and Reassessment, for the three clusters of players in the two feedback treatments.

Treatment	Stage	Centers of transitions					Signif. (Hotelling's T2)
		own-sum	own-dom	other-sum	other-dom	intracell	
				Cluster 1			
feedback/DS	A	.163	.185	.211	.353	.087	T2 = 5.72
	R	.220	.367	.130	.167	.115	Df = (4, 43)
							p < .001
feedback/DO	A	.152	.175	.211	.380	.080	T2 = 15.176
	R	.026	.204	.203	.502	.065	Df = (4, 50)
							p < .001
				Cluster 2			
feedback/DS	A	.753	.200	.010	.007	.028	T2 = 3.988
	R	.496	.229	.117	.043	.011	Df = (4, 15)
							p = .021
feedback/DO	A	.691	.229	.020	.010	.050	T2 = 9.874
	R	.098	.229	.328	.208	.136	Df = (4, 18)
							p < .001
				Cluster 3			
feedback/DS	A	.229	.090	.261	.130	.290	T2 = 10.338
	R	.279	.200	.163	.124	.234	Df = (4, 21)
							p < .001
feedback/DO	A	.205	.092	.293	.114	.296	T2 = 10.313
	R	.082	.094	.443	.166	.214	Df = (4, 9)
							p = .002

Note. In the column named "Stage", "A" indicates the Assessment, and "R" the Reassessment. Due to rounding, the reported distributions may not sum up exactly to one.

share of attention allocated to their own payoffs, and those in Cluster 3 also for that allocated to intracell comparisons.

Thus, eye-tracking data suggest the possibility that the increase of equilibrium in the *feedback/DO* treatment can mask substantially different underlying learning styles. In the Learning stage, players in the three clusters similarly learn to select the equilibrium strategy (the proportion of equilibrium choices in the last block of five games is 91%, SD = .26, for Cluster 1, 96%, SD = .13, for Cluster 2, and 85%, SD = .38, for Cluster 3; the difference across groups is not significant, F(2, 86) = 0.866, p = .424), but the learning outcomes for the three clusters are substantially different. This is confirmed by the analysis of transitions in the Learning stage (reported in the Supplementary material) and in the Reassessment. Differently than players in Cluster 1, Cluster 2 and 3 players do not learn to suppress the use of *own-sum* and *intracell* transitions to analyze the other player's payoffs, and display a large variability in the use of *other-dom* transitions (whose frequency remains lower than that for Cluster 1 players; see Table 4, Fig. 7, and Figure S.3). Thus, although paying attention to the incentives of their opponent, and partially using *other-dom* transitions, these players do not appear to carry out a systematic analysis of dominance and best-reply.

In the analysis of choice data by cluster in the Reassessment of the *feedback/DO* treatment, we showed that Cluster 2 and 3 players, although being able to learn to choose the equilibrium action in DO and ND games similarly to players in Cluster 1, fail to do so in the strategically simpler DS games. By analyzing how players in the three clusters allocate their attention to the payoffs of the opponent in the final Reassessment stage, we can see how different lookup patterns affect the frequency of equilibrium choices in DS games. In DS games, *other-dom* and *other-sum* transitions are significantly negatively correlated (the test of Pearson's product-moment correlation gives an estimate of -0.49, $t(52) = 4.107$, $p < .001$ for Cluster 1; -0.57, $t(20) = 3.064$, $p = .006$, for Cluster 2; and -0.74, $t(11) = 3.683$, $p = .003$, for Cluster 3). This means that these two types of transitions tend to be mutually exclusive: Players that make large use of one type of transitions tend not to use the other. In addition, for all three clusters, the frequency of equilibrium choices in DS games is significantly positively correlated with that of *other-dom* transitions (the test of Pearson's product-moment correlation gives an estimate of 0.44, $t(52) = 3.500$, $p < .001$ for Cluster 1; 0.60, $t(20) = 3.340$, $p = .003$, for Cluster 2; and 0.66, $t(11) = 2.931$, $p = .0137$, for Cluster 3), but negatively correlated with that of *other-sum* transitions (the test of Pearson's product-moment correlation gives an estimate of -0.58, $t(52) = 5.125$, $p < .001$ for Cluster 1; -0.88, $t(20) = 8.463$, $p < .001$, for Cluster 2; and -0.78, $t(11) = 4.209$, $p = .001$, for Cluster 3). This means that players that use *other-dom* transitions learn a best-response reasoning, and choose the equilibrium choice in DS games more frequently than players that use *other-sum* transitions, who do not appear to make a strategic analysis of their opponent's incentives. The fact that players in Cluster 1 use more often *other-dom* transitions, whereas those in Cluster 2 and 3 use more often *other-sum* ones explains the difference between the frequency of equilibrium choices in DS games illustrated above.[9] Therefore, both choice and eye tracking data illustrated earlier reject the use of best-response reasoning by players in these two clusters. Rather, Cluster 2 and 3 players are likely to use some heuristics of play that in DS games does not allow them to recognize their own dominant strategy (respectively, 30% and 52% of the time), and that systematically suggests them playing what could be labeled as the "cooperative" strategy[10] (players in Cluster 2 select this strategy 28% of the time, $SD = .43$, whereas those in Cluster 3 select it 43% of the time, $SD = .46$), although this does not necessarily imply that these are cooperative players. However, our data do not allow us to accurately reconstruct such a heuristic, and furthermore it is plausible that different players develop and apply different heuristics.

On the contrary, Cluster 1 players in the *feedback/DO* treatment learn to look for dominance in their own and in their opponent's actions, displaying a larger frequency of *other-dom* transitions, and substantially suppressing *own-sum* transitions (but not the *own-dom* ones, that are necessary for the recognition of own dominant actions and the formulation of the best-response). This reveals a deeper strategic learning compatible with the dominance analysis, which is effective in DS, DO, and ND games. As for ND games, although not solvable through iterated dominance, the combination of *own-dom*, *other-sum*, and *other-dom* transitions (*intracell* and *own-sum* transitions are almost totally suppressed) displayed by Cluster 1 players is indicative of a strategic thinking close to equilibrium reasoning also in these games, as indicated by the significant corresponding increase in the frequency of equilibrium choices (passing from 53%, $SD = .35$, to 67%, $SD = .40$; paired t-test: $t(53) = 2.820$, $p = .007$).

9 Of course, also *own-dom* transitions are positively and highly correlated with the frequency of equilibrium choices in DS games in all three clusters (the test of Pearson's product-moment correlation gives an estimate of 0.51, $t(52) = 4.253$, $p < .001$ for Cluster 1; 0.65, $t(20) = 3.873$, $p < .001$, for Cluster 2; and 0.63, $t(11) = 2.679$, $p = .0214$, for Cluster 3). However, the point here is that these transitions are not necessarily related to a more strategic approach of analysis of the game incentives, whereas the transitions used by players to analyze their counterpart's payoffs provide a more compelling elicitation of their strategic intent.

10 This strategy corresponds to Action III of the example of DS game illustrated in Fig. 1. The cooperative strategy is the row that includes the cell that yields the largest payoff sum (i.e., the cell (67, 79) in the example of Fig. 1), and in DS games it never corresponds to the equilibrium action.

It has to be noticed that in ND games, by design, the equilibrium action for the Computer is the action with the largest payoff sum and is also a "quasi-dominant" action (i.e., it is dominant but for one payoff), although it is never the action that includes the largest payoff (cf. Section S.10 of the Supplementary material). Therefore, in the feedback/DO treatment, for players in Clusters 2 and 3, the selection of the equilibrium strategy can be facilitated by the structural features of ND games just mentioned. Thus, for these players (who make substantial use of *other-sum* transitions), equilibrium choice in the ND games is likely to be supported by a heuristic that exploits some of the just mentioned features of these games. For Cluster 1 players, however, learning appears to be rather different. The large use of *other-dom* transitions does not support the use of a "largest-sum" heuristic by these players in ND games. In addition, although the use of a "quasi-dominance" heuristic is not ruled out for these players, their analysis of game payoffs reveals a reasoning that significantly differs from that of players in the other two clusters and more closely conforms to the detection of dominance and best reply, if not proper equilibrium reasoning. Figures S1 and S2 of the Supplementary material report temporal analyses of transitions (within trials and across trials) observed during the Learning stage, disaggregated by cluster. Those analyses show the different impacts that repeated play of DS and DO games with feedback has on choice behavior of players in the three clusters and further confirm the results reported in this section.

A natural question is whether experience in the feedback treatments leads to transition distributions that are different across the three clusters. To answer this question, we adopt the same maximum likelihood procedure for testing differences between independent samples of distributional changes (also referred to as "perturbations") used in section 3.2 (see details in the Supplementary material). The perturbations of transitions distributions computed for Cluster 1 and Cluster 2 players in the *feedback/DO* treatment differ significantly in the center and variance structure (we reject H_{03} in favor of H_{gen}, $X^2(4) = 13.80$, p = .008; see Supplementary material for the description of the hypothesis structure), as well as for Cluster 1 and 3 (we reject H_{03} in favor of H_{gen}, $X^2(4) = 13.43$, p = .009; see Supplementary material for the description of the hypothesis structure). Similar results hold for the *feedback/DS* treatment (for Cluster 1 vs. Cluster 2, H_{03} is rejected in favor of H_{gen}, $X^2(4) = 16.99$, p = .002; for the comparison Cluster 1 vs. Cluster 3, H_{01} is rejected in favor of H_{gen}, $X^2(10) = 183.84$, p < .001). These results provide evidence that experience produces changes in the lookup patterns that are significantly different between the more strategic players in Cluster 1 and players in the other two clusters.

5. Discussion

The level of strategic sophistication of an individual is commonly intended as his/her ability to anticipate the moves of the other decision makers in an interactive decision problem, based on the analysis of his/her own incentives as well as those of his/her counterparts. Although previous research has shown that people can be taught more sophisticated strategies of play (Zonca et al., 2019), or that cognitive skills and personal traits affect how people learn equilibrium play (Gill and Prowse, 2016), the question of how individuals endogenously extract knowledge from experience and use it to enhance their strategic skills is substantially unanswered. In this article, we test the hypothesis of whether and under what conditions experience affects players' strategic sophistication. To this end, we first assess participants' initial level of strategic sophistication by analyzing choice and eye-tracking data on three classes of 3×3 two-person strategic-form games (DS games, in which only the participant has a dominant action; DO games, in which only the opponent, i.e., the Computer, has a dominant action; and ND games, in which neither of the players has a dominant action) that were played without feedback (Assessment stage). Subsequently, we train participants on a sequence of either DS or DO games, played with or without immediate feedback, in a between-subject design (Learning stage). Finally, we reassess participants' strategic sophistication on the same three families of games as done in the first stage (Reassessment stage).

We show that strategic sophistication and equilibrium play are significantly enhanced only in presence of feedback. Furthermore, although feedback in DS and DO games provides the same opportunity of strategic learning, we observe that participants improve their strategic skills only if trained on DO games. More in detail, repeated play of games in which participants do have a dominant action but their opponent does not (i.e., DS games) prevents them from shifting their attention to their opponent's incentives, and thus does not increase their level of sophistication. In particular, the absence of a dominant action for the Computer makes the task of predicting its choices more difficult than in DO games, and could discourage players from inspecting consistently its payoffs. We observe that participants that are exposed to repeated play of DS games learn to make use of lookup patterns that are associated with the search for their own strategy with the largest payoff sum. On the contrary, participants that repeatedly play games in which they do not have a dominant strategy (i.e., DO games) do learn to allocate significant part of their attention to the other player's incentives, displaying a process of acquisition of visual information that reflects an enhanced level of strategic sophistication. These participants increase the share of lookup patterns that are associated with the search of a dominant action of the other player, and generalize such learning to recognize and select the equilibrium action also in the ND games, which cannot be solved by iterated dominance. Such transfer of strategic approach across game types is confirmed by the observation that players use the same lookup patterns while solving the three different game types (see Figure S5 and S6 of the Supplementary material).

Do all players that play DO games with feedback in the Learning stage learn in the same way? The individual analysis of choice and eye-tracking data reveals an interesting result. We classify players according to their strategic sophistication at the beginning of the experiment, before the manipulation of experience in the Learning stage has occurred, and we obtain three clusters. The data show that players that already had a more strategic approach (participants in Cluster 1) responded to repeated play with feedback of DO games in a very different way than those who initially displayed a strategically naive approach (participants in Cluster 2 and 3). Players that are initially more naive appear to learn a heuristic of play that allows them to match the equilibrium play of the more sophisticated players of Cluster 1 in DO and ND games, but not in the "strategically simpler" DS games. This result is explained by the observation that although naive players do learn to allocate attention to the payoffs of the other player, they do that mostly by adopting lookup patterns that are not compatible with the detection of dominance and, more in general, with equilibrium reasoning. After having gained experience with DO games, also sophisticated players learn to pay more attention to the payoffs of their opponent, but mostly through lookup patterns that reveal intentional search for dominance in their opponent's actions, and thus correspond to a deeper strategic analysis—if not proper equilibrium reasoning. This is reflected by the increased frequency of equilibrium choices in DS, DO, and ND games in the final stage of the experiment for these players.

A word of caution is due for what concerns the persistency of the learned strategic skills. In this paper we refer to the "newly acquired" strategic skills of participants, but our experimental design does not test to what extent the observed enhancement in strategic sophistication is retained. This important question about the temporal stability of learned strategic skills will be addressed by future work.

The milestone study by Selten et al. (2003) shows that repeated play with feedback of (randomly generated) two-person 3×3 games enhances participants' rational reasoning, significantly increasing the proportion of equilibrium choices across trials.[11] Our study enriches and qualifies Selten et al.'s (2003) conclusion, by analyzing strategic learning at a higher level of detail, which allows us to distinguish between substantially different types of learning and to draw boundary conditions for strategic learning to occur.

Our analysis of the ocular movements and of how these are linked to strategic sophistication is inherently connected with the normal form representation of our games. Thus, the extent to which our results can be generalized to games presented in different forms (e.g., in the extensive form) is an open question. However, we believe that our general considerations on the plasticity of strategic sophistication and the observation that players can learn very different choice strategies from the same type of experience have a general validity that goes beyond our specific experimental settings.

6. Conclusions

Our data provide evidence that strategic sophistication is importantly shaped by experience with previous strategic decisions, but also that the reactions to this experience can be quite heterogeneous. The construct of strategic learning goes well beyond the sole learning of sophisticated equilibrium thinking, but also encompasses the development of simple choice rules that were reinforcing in similar situations in the past. This generalization hypothesis has been theorized and empirically confirmed in psychology studies of individual decision making (see, for example, Gonzalez and Dutt, 2011; Plonsky et al., 2015; Marchiori et al., 2015), but also in economics studies of similarity across games (see, for example, Gilboa and Schmeidler, 1995; Knez and Camerer, 2000; Devetag, 2005; Mengel, 2012; Grimm and Mengel, 2012). In addition, our results further confirm that equilibrium play can mask different underlying strategic intents and types of strategic learning, not necessarily consistent with equilibrium reasoning (see also Crawford et al., 2013). Even in those situations in which incentives are unambiguously and fully described to players and the behavior of the strategic counterpart is unambiguously identifiable (as in our experiment), experience can affect decision makers' behavior in radically, when not unexpectedly, different ways.

Funding sources

This work has been funded by grants from the Danish Council for Independent Research (DFF – 7015 – 00050, Davide Marchiori; DFF – 0166 – 00005B, Sibilla Di Guida). The funding sources were not involved in the study design; data collection; analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Acknowledgments

This paper greatly benefited from discussions with Rosemarie Nagel and Jens Schmidt on a related project. We thank participants of the "Society for NeuroEconomics Conference 2019"

[11] This study was conducted adopting the strategy method for eliciting players' reasoning, and the feedback consisted in the performance of a given strategy matched against all other submitted strategies in a computerized tournament.

and the "SABE IAREP 2019", and two anonymous reviewers for their insightful suggestions and comments. We gratefully acknowledge the valuable help of Giorgio Marinato (MoMiLab, IMT School for Advanced Studies Lucca) in the data collection and of Massimo Vescovi (CIMEC, University of Trento) for granting us access to the lab facilities.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2021.105291.

References

Agranov, M., Potamites, E., Schotter, A., Tergiman, C., 2012. Beliefs and endogenous cognitive levels: an experimental study. Games Econ. Behav. 75 (2), 449–463.

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman & Hall, London.

Alaoui, L., Penta, A., 2015. Endogenous depth of reasoning. Rev. Econ. Stud. 83 (4), 1297–1333.

Arieli, A., Ben-Ami, Y., Rubinstein, A., 2011. Tracking decision makers under uncertainty. Am. Econ. J. Microecon. 3 (4), 68–76.

Aumann, R., Brandenburger, A., 1995. Epistemic conditions for Nash equilibrium. Econometrica 63 (5), 1161.

Bayer, R.C., Renou, L., 2016. Logical abilities and behavior in strategic-form games. J. Econ. Psychol. 56, 39–59.

Brandenburger, A., 1992. Knowledge and equilibrium in games. J. Econ. Perspect. 6 (4), 83–101.

Brocas, I., Carrillo, J.D., Sachdeva, A., 2018. The path to equilibrium in sequential and simultaneous games: a mouse-tracking study. J. Econ. Theory 178, 246–274.

Brocas, I., Carrillo, J.D., Wang, S.W., Camerer, C.F., 2014. Imperfect choice or imperfect attention? Understanding strategic thinking in private information games. Rev. Econ. Stud. 81 (3), 944–970.

Camerer, C., Ho, T.-H., 1999. Experience-weighted attraction learning in normal form games. Econometrica 67 (4), 827–874.

Carpenter, J., Graham, M., Wolf, J., 2013. Cognitive ability and strategic sophistication. Games Econ. Behav. 80, 115–130.

Chen, C.T., Huang, C.Y., Wang, J.T.Y., 2018. A window of cognition: eyetracking the reasoning process in spatial beauty contest games. Games Econ. Behav. 111, 143–158.

Costa-Gomes, M.A., Crawford, V.P., 2006. Cognition and behavior in two-person guessing games: an experimental study. Am. Econ. Rev. 96 (5), 1737–1768.

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

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.

Devetag, G., 2005. Precedent transfer in coordination games: an experiment. Econ. Lett. 89 (2), 227–232.

Devetag, G., Warglien, M., 2003. Games and phone numbers: do short-term memory bounds affect strategic behavior? J. Econ. Psychol. 24 (2), 189–202.

Devetag, G., Di Guida, S., Polonio, L., 2016. An eye-tracking study of feature-based choice in one-shot games. Exp. Econ. 19 (1), 177–201.

Erev, I., Roth, A.E., 1998. Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. Am. Econ. Rev. 88 (4), 848–881.

Fehr, D., Huck, S., 2016. Who knows it is a game? On strategic awareness and cognitive ability. Exp. Econ. 19 (4), 713–726.

Fraley, C., Raftery, A.E., 2002. Model-based clustering, discriminant analysis, and density estimation. J. Am. Stat. Assoc. 97 (458), 611–631.

Franco-Watkins, A.M., Johnson, J.G., 2011. Applying the decision moving window to risky choice: comparison of eye-tracking and mousetracing methods. Judgm. Decis. Mak. 6 (8), 740.

Georganas, S., Healy, P.J., Weber, R.A., 2015. On the persistence of strategic sophistication. J. Econ. Theory 159, 369–400.

Gilboa, I., Schmeidler, D., 1995. Case-based decision theory. Q. J. Econ. 110 (3), 605–639.

Gill, D., Prowse, V., 2016. Cognitive ability, character skills, and learning to play equilibrium: a level-k analysis. J. Polit. Econ. 124 (6), 1619–1676.
Glöckner, A., Herbold, A.K., 2011. An eye-tracking study on information processing in risky decisions: evidence for compensatory strategies based on automatic processes. J. Behav. Decis. Mak. 24 (1), 71–98.
Gonzalez, C., Dutt, V., 2011. Instance-based learning: integrating sampling and repeated decisions from experience. Psychol. Rev. 118 (4), 523.
Graffeo, M., Polonio, L., Bonini, N., 2015. Individual differences in competent consumer choice: the role of cognitive reflection and numeracy skills. Front. Psychol. 6, 844.
Grimm, V., Mengel, F., 2012. An experiment on learning in a multiple games environment. J. Econ. Theory 147 (6), 2220–2259.
Knez, M., Camerer, C., 2000. Increasing cooperation in prisoner's dilemmas by establishing a precedent of efficiency in coordination games. Organ. Behav. Hum. Decis. Process. 82 (2), 194–216.
Knoepfle, D.T., Wang, J.T.Y., Camerer, C.F., 2009. Studying learning in games using eye-tracking. J. Eur. Econ. Assoc. 7 (2–3), 388–398.
Marchiori, D., Di Guida, S., Erev, I., 2015. Noisy retrieval models of over- and under-sensitivity to rare events. Decision 2 (2), 82.
Marchiori, D., Warglien, M., 2008. Predicting human interactive learning by regret-driven neural networks. Science 319 (5866), 1111–1113.
Mauersberger, F., Nagel, R., 2018. Levels of reasoning in Keynesian Beauty Contests: a generative framework. In: Handbook of Computational Economics, vol. 4. Elsevier, pp. 541–634.
Mengel, F., 2012. Learning across games. Games Econ. Behav. 74 (2), 601–619.
Nagel, R., 1995. Unraveling in guessing games: an experimental study. Am. Econ. Rev. 85 (5), 1313–1326.
Pawlowsky-Glahn, V., Egozcue, J.J., Tolosana-Delgado, R., 2015. Modeling and Analysis of Compositional Data. John Wiley & Sons.
Plonsky, O., Teodorescu, K., Erev, I., 2015. Reliance on small samples, the wavy recency effect, and similarity-based learning. Psychol. Rev. 122 (4), 621.
Polonio, L., Coricelli, G., 2019. Testing the level of consistency between choices and beliefs in games using eye-tracking. Games Econ. Behav. 113, 566–586.
Polonio, L., Di Guida, S., Coricelli, G., 2015. Strategic sophistication and attention in games: an eye-tracking study. Games Econ. Behav. 94, 80–96.
Proto, E., Rustichini, A., Sofianos, A., 2019. Intelligence, personality, and gains from cooperation in repeated interactions. J. Polit. Econ. 127 (3), 1351–1390.
Selten, R., Abbink, K., Buchta, J., Sadrieh, A., 2003. How to play (3×3)-games: a strategy method experiment. Games Econ. Behav. 45 (1), 19–37.
Slonim, R.L., 2005. Competing against experienced and inexperienced players. Exp. Econ. 8 (1), 55–75.
Wang, J.T.Y., Spezio, M., Camerer, C.F., 2010. Pinocchio's pupil: using eye tracking and pupil dilation to understand truth telling and deception in sender-receiver games. Am. Econ. Rev. 100 (3), 984–1007.
Zonca, J., Coricelli, G., Polonio, L., 2019. Does exposure to alternative decision rules change gaze patterns and behavioral strategies in games? J. Econ. Sci. Assoc. 5 (1), 14–25.
Zonca, J., Coricelli, G., Polonio, L., 2020a. Gaze data reveal individual differences in relational representation processes. J. Exp. Psychol. Learn. Mem. Cogn. 46 (2), 257–279.
Zonca, J., Coricelli, G., Polonio, L., 2020b. Gaze patterns disclose the link between cognitive reflection and sophistication in strategic interaction. Judgm. Decis. Mak. 15 (2), 230–245.