Tulane Economics Working Paper Series

Social Metacognition: A Correlational Device for Strategic Interactions

Chiara Scarampi
University of Geneva
chiara.scarampi@unige.ch

Richard Fairchild
University of Bath
r.fairchild@bath.ac.uk

Luca Fumarco
Tulane University
lfumarco@tulane.edu

Alberto Palermo
Trier University
palermo@iaaeu.de

Neal Hinvest
University of Bath
n.hinvest@bath.ac.uk

Working Paper 2111
June 2021

Abstract

This study reports a laboratory experiment wherein we investigate the role of social metacognition– i.e., the ability to monitor and control one's own and others' mental states – in a chicken game. In the first part of the experiment, we try to implement a correlated equilibrium, a generalisation of the Nash equilibrium where players' strategies are correlated by a third party/mechanism/choreographer. We find that social metacognition is a signif- icant predictor of subjects' strategy choices. The experiment proceeds without third party recommendations. We find evidence that subjects with high social metacognition are more likely to play a correlated equilibrium; that is, social metacognition acts "as if" it is the correlating mechanism. We relate our findings to the individual social metacognitive ability as well as to the group composition.

Keywords: Correlated Equilibrium, Social Metacognition, Experimental Economics
JEL codes: C72, C92, D91

Social Metacognition: A Correlational Device for Strategic Interactions*

Chiara Scarampi†, Richard Fairchild‡, Luca Fumarco§, Alberto Palermo¶, Neal Hinvest‖

Abstract

This study reports a laboratory experiment wherein we investigate the role of social metacognition– i.e., the ability to monitor and control one's own and others' mental states – in a chicken game. In the first part of the experiment, we try to implement a correlated equilibrium, a generalisation of the Nash equilibrium where players' strategies are correlated by a third party/mechanism/choreographer. We find that social metacognition is a significant predictor of subjects' strategy choices. The experiment proceeds without third party recommendations. We find evidence that subjects with high social metacognition are more likely to play a correlated equilibrium; that is, social metacognition acts "as if" it is the correlating mechanism. We relate our findings to the individual social metacognitive ability as well as to the group composition.

Keywords: Correlated Equilibrium, Social Metacognition, Experimental Economics

JEL codes: C72, C92, D91

*We are grateful to John Duffy and Nick Feltovich for sharing the code to run their experiment on Z-Tree. We thank John Duffy and David Levine for valuable comments that helped improve the paper. We also thank participants at several conferences for helpful suggestions. All remaining errors are ours.
This work was supported by the University of Bath under a Research Studentship Award and a Psychology Department Postgraduate Support Fund awarded to Chiara Scarampi. Declarations of interest: none.

†Corresponding author: Chiara Scarampi, Swiss National Center of Competences in Research LIVES – Overcoming Vulnerability: Life Course Perspectives, Geneva and Lausanne, Switzerland and Centre for the Interdisciplinary Study of Gerontology and Vulnerability (CIGEV), University of Geneva, Boulevard du Pont-d'Arve 28, 1205 Geneva, Switzerland. E-mail: chiara.scarampi@unige.ch

‡School of Management, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom. E-mail: r.fairchild@bath.ac.uk

§Department of Economics and the Murphy Institute, Tulane University, 206 Tilton Hall 6823 St. Charles Avenue New Orleans, Louisiana 70118. E-mail: lfumarco@tulane.edu

¶Institute for Labour Law and Industrial Relations in the European Union (IAAEU), Trier University, Behringstr. 21, 54296 Trier, Germany. E-mail: palermo@iaaeu.de

‖Department of Psychology, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom. E-mail: n.hinvest@bath.ac.uk

1 Introduction

> Human behaviour is most fruitfully modelled as the interaction of rational agents with a social epistemology, in the context of social norms that act as correlating devices that *choreograph* social interaction (Gintis, 2009, p. xiii).

The most used equilibrium concept in non-cooperative game theory is the Nash equilibrium, which is theoretically found assuming that players' strategies are *uncorrelated*. The alternative equilibrium concept is that of correlated equilibrium (Aumann, 1974), which represents a generalisation of the Nash equilibrium. The correlated equilibrium departs from the Nash equilibrium in assuming the presence of a third party, mechanism, or *choreographer* as named by Gintis (2009), which correlates the strategies of the players.[1] In this paper, we aim at investigating whether social metacognition – i.e., the ability to reflect upon, monitor and control one's own and others' knowledge, emotions and actions (Jost et al., 1998) - can help individuals correlate their behaviour in strategic interactions *without* requiring the presence of an external correlating device.

Broadly, a correlated equilibrium is a known distribution of signals over the set of possible "scenarios" in which players can find themselves and it requires a recommendation about the strategy to play given to players by a choreographer. Aumann (1974) showed that in the set of possible correlated equilibria there could be some that are payoff-enhancing with respect to the Nash equilibrium in mixed strategies. However, despite this appealing result, the correlated equilibrium is not so often invoked. It appears that researchers find the correlated equilibrium hardly justifiable as solution concept due to the requirement of a third party/mechanisms or even a possible communication among players. Nevertheless, it is famous Roger Myerson's claim: "If there is intelligent life on other planets, in a majority of them, they would have discovered correlated equilibrium before Nash equilibrium". More recently, the use of the correlated equilibrium has attracted a deeper interest especially in experimental economics. Our analysis appears relevant because metacognition is a teachable skill, (Schraw, 2001; Lecce et al., 2015).

[1] Myerson (1991) uses the term *mediator* instead of *choreographer*.

The rationale for this research relies on the relation between *social metacognition* and the *common knowledge* postulate, a critical element to predict behaviour in strategic interaction games. The common knowledge postulate is a recursive argument: knowledge of others' knowledge, knowledge of their knowledge of one's knowledge, ad infinitum "common knowledge". Aumann (1976) argued that individuals reflect on the intentions and actions of the others they are interacting with and know that they are doing the same. Common knowledge has been considered a recursive form of mentalising (De Freitas et al., 2019). Similarly, but less narrowly, social metacognition is the ability to infer others' beliefs, intentions and behaviours. In this paper, we hypothesise that this cognitive mechanism can help individuals with high social metacognition (interacting with other highly metacognitive individuals) to correlate their strategies because they realise a common knowledge of their ability.

We conduct an experiment in two phases where players are faced with a Chicken game. In the first phase, we follow Duffy and Feltovich (2010) and provide subjects with recommendations. In particular, the suggested correlated equilibrium is payoff-enhancing compared to the Nash equilibrium in mixed strategies. In the second phase, we leave subjects playing the same game without recommendations. We account for individual differences by measuring individual levels of social metacognition.

For the rounds with recommendations, first, we replicate the findings of Duffy and Feltovich (2010). Intuitively, since the suggested correlated equilibrium is payoff-enhancing compared to the available Nash equilibria in mixed strategy, subjects follow the recommendations. Put differently, subjects can realise the beneficial (also individual) effect of correlating strategies. Second, we find that social metacognition is a predictor of which strategy players adopt. Hence, we conduct analyses hypothesising that subjects with a higher social metacognition adapt their choices according to group composition.

After the rounds with recommendations, subjects are left playing without recommendations. We hypothesise, and provide evidence, that subjects with higher social metacognitive abilities will be able to better coordinate their behaviour, and their play, over time, will converge towards a correlated equilibrium.

The paper proceeds as follows. The next section discusses the relevant literature, providing a deeper look at the concept of social metacognition and its link with previous economics studies which (in)directly referred to it. In section 3, we describe the experiment and the method we used to measure social metacognition. Section 4 presents the results and section 5 offers a discussion of the main findings. The appendix contains the questionnaire used to asses the level of social metacognition and reports the procedure and statistical tests confirming the validity of the instrument.

2 Related literature

Human beings are characterised by the ability to form knowledge, understanding and sharing of mental states, which in psychology is referred to as social metacognition (Frith, 2012; Jost et al., 1998). Through monitoring and control processes, social metacognition allows individuals to create an accurate representation of the mental states of others (also known as mentalising or theory of mind) (Efklides, 2014; Frith and Frith, 1999; Frith, 2012) and use this information to predict others' intentions and determine the most appropriate behaviour for the specific situation (Efklides, 2008). As Volet et al. (2009a) pointed out, it involves the application of skills and strategies learned through instruction and previous social interactions to the ongoing interactive or collaborative context, informs on the strategies to use, and contributes to meaning making and regulation of one's own cognitive performance in social contexts.

Recent studied have provided some evidence that the ability to infer the mental states of the opponents can lead to better outcomes. That is, social metacognition allows individuals to apply to novel contexts information learnt in prior situations, resulting in the ability to anticipate the opponents' intentions and behaviour, and optimally respond to their strategies (Camerer et al., 2005; Robalino and Robson, 2016). Fehr and Huck (2016) examined indirectly the relationship between economic behaviour and social metacognition by looking at "strategic awareness" - i.e., the awareness of the importance of considering the abilities of the opponents when reasoning about how to play the game. The authors elicited measures of subjects' cognitive abilities and beliefs about others' cognitive abilities in Nagel (1995)'s classic Beauty Contest Game, where

participants were asked to pick a number between 0 and 100. The results showed that participants with high cognitive abilities avoided numbers above 50 and their choices correlated with their beliefs about others' cognitive abilities. In contrast, choices made by participants with low cognitive abilities were randomly distributed over the whole interval and did not correlate with their beliefs about others' cognitive abilities. This suggests that the use of social metacognition to reason about others is essential for playing the game. In the current study, we add to this literature by providing evidence that social metacognition implies different behaviours to different opponents.

Further evidence on the relevance of social metacognition for strategic interactions comes from a series of neurocognitive studies investigating the neural correlates of strategic interactions and finding activations in a set of brain regions which are consistently associated with mentalising (e.g., Powell et al., 2017).[2] In particular, these activations are stronger or found only when subjects play against another (alleged) human and not when they play against a computer (McCabe et al., 2001; Gallagher et al., 2002; Rilling et al., 2004). These results suggest that strategic interactions with other human beings are not based merely on computational analyses, but require also the use of social metacognitive processes to reason about the mental states of the opponents and decide accordingly.

Shteynberg et al. (2020) have recently proposed a theory of collective learning in which they discuss a form of collective attention which represents shared subjective states as common knowledge. This metacognitive capacity yields mutually known representations, emotions, evaluations, and beliefs which facilitate cognitive alignment among group members, enhancing in turn social coordination. Some preliminary evidence of the connection between social metacognition and the ability to solve interpersonal coordination games comes from Curry and Chesters (2012). They studied couples of participants who were presented with some numerical-verbal and visual-spatial coordination problems and asked to try and give the same answer as their co-player. The results pointed to an association between self-report scores of mentalising abilities and coordination success.

[2]See Frith and Frith (1999, 2006) for a review of the neurobiology of mentalising.

The research question that we want to address is then linked to an observation made by De Freitas et al. (2019). The authors claimed that most research has treated the processes used by individuals to recognise and represent common knowledge as a black box. With this research, we attempt to take a first look inside the black box by studying how social metacognition can be used to coordinate with others on a correlated equilibrium. To the best of our knowledge, no previous studies have investigated how social metacognition and the ability to take into account mental states of others relate to the correlated equilibrium. In a theoretical set-up, Foster and Vohra (1997) were the first to show that if players use a learning rule which is a calibrated forecasts of other opponents' future strategy then the resulting long-run strategy profile is in the set of correlated equilibria. Yet, Hart and Mas-Colell (2000) also showed that in a N-person game there exists a simple adaptive procedure which generates convergence to the set of correlated equilibria.[3] Arifovic et al. (2019) used some simulations to test an algorithm providing evidence of convergence to the correlated distribution. The authors showed that the algorithm does not require knowledge of the distribution of recommendations to learn the correlated equilibrium. Most of the empirical literature on the correlated equilibrium has focused on whether participants tend to follow or ignore recommendations given by a third party. Closely related to our study are Cason and Sharma (2007) and Duffy and Feltovich (2010), from which we depart adding an analysis to disentangle the role of social metacognition in coordination games.

Cason and Sharma (2007) attempted to implement a correlated equilibrium with payoffs outside the convex hull of Nash equilibrium payoffs by privately recommending strategies in a Chicken game. Their results suggest that individuals do not follow recommendations inducing correlated equilibria. However, more experienced participants in the experiment tended to follow recommendations more frequently in the second half of trials, suggesting therefore that after a learning process subjects could start following recommendations. Speculating on the reason why subjects do not follow recommendations from the start – and in line with our theoretical explanation based on social metacognition – the authors hypothesised that individuals may form beliefs about their opponents' mistakes and update them in subsequent periods. As a possible

[3]See Hart and Mas-Colell (2000) and references therein also for alternative procedures converging to the set of correlated equilibria.

further research avenue, the authors stressed the relevance of providing a theoretical argument able to explain what conditions would lead individuals to build more and more accurate beliefs about the probability assigned to opponents following recommendations.

Conversely, Duffy and Feltovich (2010) used a Chicken game to explore the empirical validity of the correlated equilibrium with third-party recommendations drawn from different publicly announced distributions. They found that individuals do not blindly follow recommendations and the likelihood to follow a recommendation depends on the underlying distribution of outcomes. More precisely, a correlated equilibrium that is payoff-enhancing relative to the available Nash equilibria in mixed strategy is a necessary condition for recommendations to have any substantial effect on behaviour.

In this paper, we examine social metacognition as a cognitive process underlying coordination. More precisely, we want to observe if social metacognition – as ability to interconnect people's thinking and deriving actions – can help people coordinate their decisions and allow them to obtain better results out of their interactions; i.e., social metacognition works "as if" it is *choreographing* people.

3 Experiment and Measures

3.1 Participants

A total of 98 subjects (age range 21-79, $M = 43.67$ years, $SD = 19.76$; 55 female) participated in the study, which was conducted at the University of Bath, UK. Participants were recruited online and in the community with advertisements in newspapers, forums, newsletters and social media. Participants received a £5 show-up fee plus up to £9 depending on their performance (see the Procedure section below). All participants were healthy and free from neurological and psychiatric disease. They gave their consent to participate in the study and the research was approved by the University of Bath Psychology Ethics Committee.

3.2 The game

As in Duffy and Feltovich (2010), we designed and conducted an experiment in which subjects played the Chicken game shown in Figure 1, which has correlated equilibria with payoffs that lie outside the convex hull of Nash equilibrium payoff pair.

		Player 2	
		D	C
Player 1	D	0,0	9,3
	C	3,9	7,7

Figure 1: The basic Chicken game

This game has two Nash equilibria in pure strategies (C, D) and (D, C) and one equilibrium in mixed strategies where the action C is selected with probability 0.6. The payoffs associated with these three equilibria are $(3, 9)$, $(9, 3)$, and $(5.4, 5.4)$ respectively. However, in this game players can do even better by using third party recommendations. Suppose the distribution of recommended strategy profiles is that in Figure 2.

		Player 2	
		D	C
Player 1	D	0	1/3
	C	1/3	1/3

Figure 2: Recommended play

Additionally, suppose that players are given only their suggested strategy. Using this (correlated equilibrium) distribution over the possible outcomes of the game, the expected payoff is $19/3 > 5.4$. That is, the proposed correlated equilibrium is payoff-enhancing compared with the Nash equilibrium in mixed strategies.

The set of correlated equilibria is found as follows. Define the probabilities of the outcomes (C, C), (C, D), (D, C), and (D, D) as α, β, γ, and δ respectively. Each player should maximise their expected payoff given the signal (recommendation) they receive. If Player 1 is recommended to play C, then the conditional probability that the chosen outcome is (C, C) is $\alpha/(\alpha+\beta)$, whereas

the conditional probability that the chosen outcome is (C, D) is $\beta/(\alpha+\beta)$. Under the belief that Player 2 will follow the received recommendation, Player 1's conditional expected payoff is $7\alpha/(\alpha+\beta) + 3\beta/(\alpha+\beta) = (7\alpha+3\beta)/(\alpha+\beta)$ from following the C recommendation and $9\alpha/(\alpha+\beta) + 0\beta/(\alpha+\beta) = 9\alpha/(\alpha+\beta)$ from ignoring the recommendation and choosing D. Player 1 will then prefer to follow the C recommendation if $(7\alpha+3\beta)/(\alpha+\beta) \geqslant 9\alpha/(\alpha+\beta)$, that is $3\beta \geqslant 2\alpha$.

With a similar reasoning, we obtain the four conditions for a correlated equilibrium:

- Player 1 will prefer to follow the C recommendation if $3\beta \geqslant 2\alpha$;
- Player 1 will prefer to follow the D recommendation if $2\gamma \geqslant 3\delta$;
- Player 2 will prefer to follow the C recommendation if $3\gamma \geqslant 2\alpha$;
- Player 2 will prefer to follow the D recommendation if $2\beta \geqslant 3\delta$.

A correlated equilibrium is the distribution of the probabilities α, β, γ, and δ which satisfies the four inequalities above and $\alpha+\beta+\gamma+\delta = 1$.

3.3 Social metacognition measure

Social metacognition was measured through the administration of a new questionnaire: the Social Metacognition Inventory (SMI; see the supplementary materials for a detailed list of the items and information about the instrument validation). Respondents were asked to rate each item of the questionnaire on a 7-point Likert scale (*1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = neither disagree nor agree, 5 = slightly agree, 6 = agree, 7 = strongly agree*).

3.4 Classification of groups based on social metacognition

Since the correlated equilibrium is based on coordination amongst individuals, we believe that strategy choices in the Chicken game depend not only on an individual's ability to reflect on others' mental states but also on the social metacognition of the opponents. Accordingly, some of the analyses presented in section 4 below compare individual behaviour depending on the level of social metacognition in the group they are playing in. More precisely, we use four methods to classify the groups based on the social metacognition scores obtained at the SMI. In the first

method, we compute the mean level of social metacognition in each group and classify as "low metacognition" the groups where the mean score at the SMI is smaller than the sample mean and "high metacognition" the groups where the mean score at the SMI is greater than the sample mean (Method 1). Since the distribution of social metacognition in the sample is left-skewed, we perform the same analysis also operating a median split; i.e., defined by the comparison between the group-wise average social metacognition and the sample median (Method 2). This allows us to conduct an additional robustness check. In order to investigate possible non-linear effects of social metacognition, we also divide the sample in quartiles (Method 3). Finally, we reason that results may be different if we considered the number of highly metacognitive subjects in the group, instead of the average metacognition level in the group. In the latter, the division of the sample in two subsamples can be driven by few subjects with very high scores at the SMI.

We hypothesise that group composition, as for the number of subjects with relatively high social metacognition (i.e., high score at the SMI), influences individual choices and then results at the group level. On one hand, if there are many subjects with a high social metacognition in a group, highly metacognitive individuals may attempt and successfully correlate their strategy in the game. On the other hand, with the presence of only a few subjects with a high social metacognition, highly metacognitive individuals may realise that attempting coordination with the other players is not feasible because opponents could fail to account for a possible payoff-enhancing randomisation. In order to study if strategies change according to group composition, we classify as "low metacognition" groups where 50% of subjects or more have a metacognitive score lower than the sample mean and as "high metacognition" groups where fewer than 50% of subjects have a score lower than the sample mean (Method 4).

Similarly, in the aggregate analyses reported in section 4 below, we use as regressors the two measures of social metacognition at the group level: (i) the mean score at the SMI in a group, and (ii) the proportion of subjects in a group whose social metacognition is higher than the sample median.

3.5 Procedure

Participants were randomly assigned to 13 groups – 8 groups of 6 subjects and 5 groups of 10 subjects – and no one played in more than one group. Each group interacted for 40 rounds, the first 20 with recommendation and the last 20 without recommendation. Duffy and Feltovich (2010) varied the order of the rounds so that in half of the sessions, the rounds without recommendations came first, and in the other half, participants started from the rounds with recommendations. They found that the order of recommendation and non-recommendation rounds did not affect subjects' choices. Based on this well-known result, we did not vary the (non)recommendation order; in our experiment, all the subjects started with 20 rounds with recommendations. Duffy and Feltovich (2010) did not account for individuals' metacognition differences, whereas we did account for them. With our experimental design, all the subjects were exposed to a learning process and realised the correlation mechanism. This allowed us to disentangle the effect of metacognition on learning.

Before starting the experiment, participants were given a consent form and a set of written instructions. After reading the instructions and signing the consent form, subjects were asked to solve a quiz to assess their understanding of the instructions. Each quiz was then graded by the experimenter and any incorrect answers were discussed.

In line with Duffy and Feltovich (2010), we used a neutral terminology in the instructions, referring to partner/opponent as "the player you are matched with". Furthermore, we did not force participants to follow the recommendations. Players were instructed about the outcome probability distribution and the recommendations given during the game only conveyed information about their part of the recommended strategy profile and not the other players' part. Only after choosing the preferred action, they were shown the recommendation received by the opponent. Furthermore, participants were not given information about the results of any other pairs of subjects, either individually or in aggregate.

The experiment was run with the software z-Tree (Fischbacher, 2007) on networked computers in a Psychology laboratory consisting of multiple individual testing booths. To avoid incentives for reputation, participants were randomly paired, according to a round-robin matching format.

They were not allowed to communicate with each other and were not given identifying information about their opponents in any round.

A round of the game *with* recommendations (rounds 1 to 20) began by showing participants their recommended action, which was randomly drawn from the appropriate aforementioned outcome distribution. Then, they were asked to choose one of the two available actions. After all participants made their decision, each subject was shown the following information: own recommendation, own choice, opponent recommendation, opponent choice, own payoff, and opponent's payoff. After observing the results, participants were redirected to the following round. In a round of the game *without* recommendations, the sequence of play was the same except for the recommendations.

At the end of round 40 (i.e., the very last round), participants were asked to answer some questions to describe what their intentions and strategies were during the game with and without recommendation. They then filled in the Social Metacognition Inventory and answered a few demographic questions.

At the end of the experiment, one of the 20 rounds with recommendations and one of the 20 rounds without recommendations were randomly chosen. Each subject received their earnings from these two rounds, at an exchange rate of £0.50 per point, together with a £5 show-up fee. The experiment typically lasted 60 minutes and total earnings per participant averaged about £10.

4 Results

Each subject played 40 rounds (20 rounds with recommendations and 20 rounds without recommendations), giving us a total of 3,920 observations, 1,960 for each of the two conditions. In what follows, we first illustrate behaviour at the individual level in the game with recommendations and then present the main results for the game without recommendations. All data analyses were conducted in Stata version 16.

4.1 The effects of recommendations and social metacognition on individual choices

First, we examine how participants treat the particular recommendations they receive. We perform a series of regression analyses using a linear probability model, which allows for direct interpretation of the coefficients as probabilities. The dependent variable is subject's choice of action - more precisely, it is an indicator of a C choice. Table 1 shows three alternative model specifications, differing in which explanatory variables are included. In the first model specification, the independent variables are the score of social metacognition obtained by the participants at the SMI and an indicator for the given recommendation (namely a C recommendation with value 1 and a D recommendation with value 0). Additionally, we include a variable for the round number and the product of the given recommendation with the round number, with the aim of capturing any existing time-varying effects of recommendations. The second model specification includes age, gender, and number of participants in the group. The last model includes individual fixed effects; so, gender, age and group size variables are eliminated. Therefore, in these latter model, we benefit from the panel data structure of the dataset, and rely on both between and within individual variation. This feature allows us to control for individual unobservable characteristics. Standard errors are corrected for heteroscedasticity and adjusted for clustering by group.

The C recommendation indicator is positive and highly significant in each model specification, suggesting that subjects follow the recommendations. Models 1 and 2 show that as participants move by one round towards the end of the game, the probability of choosing C decreases by 0.4% (*Round number*). The coefficient of social metacognition (*Subject SM*) is nonsignificant. Conversely, in the model with individual fixed effects (Model 3), round number is not significant and social metacognition is a significant and negative predictor of a C choice. Hence, there is a correlation between subjects' social metacognition and a randomisation over the two strategies. Such correlation may indicate that subjects with higher social metacognition are either less prone to a cooperative behaviour or they realise that the suggested randomisation can lead to higher (expected) payoff. We hypothesise that this last aspect depends on the "opponents" and therefore

Table 1: Results of regression model. Dependent variable: Strategy C chosen in each round

	(1)	(2)	(3)
Recommendation	0.239***	0.239***	0.256***
	(0.031)	(0.032)	(0.029)
Round number	-0.004*	-0.004*	-0.004
	(0.002)	(0.002)	(0.002)
Recommendation x round	-0.001	-0.001	-0.002
number	(0.003)	(0.003)	(0.003)
Subject SM	-0.001	-0.001	-0.035***
	(0.001)	(0.001)	(0.000)
Age		-0.000	
		(0.002)	
Gender (Male)		0.013	
		(0.055)	
Session size (10		0.000	
subjects)		(0.008)	
Frequency Strategy C	0.681	0.681	0.681
Subject fixed effects	no	no	yes
Observations	1960	1960	1960
R-squared	0.056	0.056	0.260

SM stands for social metacognition.
Standard errors clustered on session in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

we proceed with additional analyses.

We then perform additional analyses using the same specification as in Model 3 of Table 1 and subsampling the data according to the different measures of social metacognition at group level described in section 3.3. The results are reported in Table 2. Models 1 and 2 report estimates obtained for groups with low or high social metacognition according to the average score in the group compared to the average score in the full sample. Models 3 and 4 report estimates obtained when we compare the average score in the group to the median score in the sample. Model 5 reports estimates from an expanded specification model where we split scores at the SMI in quartiles. The last two models report the results relative to subjects in groups where 50% of subjects or more have a metacognitive score lower than the sample mean (Model 6) and subjects in groups where fewer than 50% of subjects have a metacognitive score lower than the sample mean (Model 7). That is, Model 6 (Model 7) presents estimates for groups where relatively many subjects had a low (high) level of social metacognition.

Table 2: Results of regression model. Dependent variable: Strategy C chosen in each round according to the SM at group level

	Low group SM wrt mean (1)	High group SM wrt mean (2)	Low group SM wrt median (3)	High group SM wrt median (4)	Entire sample (5)	Few subjects with high SM wrt mean (6)	Many subjects with high SM wrt mean (7)
Recommendation	0.274***	0.240***	0.254***	0.263**	0.256***	0.266***	0.244***
	(0.041)	(0.042)	(0.033)	(0.065)	(0.029)	(0.038)	(0.051)
Round number	-0.003	-0.005	-0.005*	-0.001	-0.004	-0.002	-0.007*
	(0.003)	(0.004)	(0.003)	(0.006)	(0.002)	(0.004)	(0.003)
Recommendation x	-0.002	-0.001	-0.000	-0.005	-0.002	-0.003	-0.000
round number	(0.005)	(0.004)	(0.004)	(0.006)	(0.003)	(0.006)	(0.003)
Subject SM	-0.022***	-0.006***	-0.030***	-0.042***		-0.022***	-0.006***
	(0.001)	(0.000)	(0.001)	(0.001)		(0.001)	(0.000)
2^{nd} quartile SM scores					-0.585***		
					(0.006)		
3^{rd} quartile SM scores					-0.735***		
					(0.005)		
4^{th} quartile SM scores					-0.148***		
					(0.003)		
Frequency Strategy C	0.678	0.684	0.678	0.687	0.681	0.681	0.680
Subject fixed effects	yes	yes	yes	yes	yes	yes	yes
Observations	960	1000	1320	640	1960	1080	880
R-squared	0.290	0.231	0.275	0.231	0.260	0.226	0.304

SM stands for social metacognition.

Standard errors clustered on session in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results for the effect of recommendations and social metacognition are consistent with the results previously discussed. The (negative) correlation is, however, only an indicator that social metacognition matters in the decision. The negative effect can be theoretically explained by the fact that the strategy profile (C, C) is not an equilibrium of the game. More clearly, subjects with high social metacognition easily realise it and rather randomise. Nevertheless, there is here evidence that these individuals adopt different strategies according to the group composition. Although negative and highly significant in each model specification, the magnitude of the effect of social metacognition varies according to the method used to classify groups as high or low on social metacognition.

The results of Model 5 suggest a convex relationship between subjects' social metacognition and the decision to play a C strategy. On average, the effect of social metacognition on the probability of choosing a C strategy is always negative; however, its magnitude increases in the second and third quartile and, eventually, decreases in the fourth quartile. The average magnitude of the effect of social metacognition for the second, third, and fourth quartile is much larger than the magnitude of the effect of social metacognition in any of the previous estimates, suggesting that metacognitive subjects in metacognitive groups are less likely to play the C strategy. Here, it is important to note that the coefficients for social metacognition in Model 5 are not immediately comparable to the coefficients in the previous model specifications. In the previous model specifications, the coefficients represent the effect of increasing subjects' social metacognition by one unit, whereas in Model 5 the coefficients represent the effect of passing from one quartile to the next one with the first quartile being the reference group.

The results of Models 6 and 7 are comparable to those of Models 1 and 2, suggesting that the methods used to classify groups based on social metacognition are robust. A one-point increase in the score of social metacognition for a subject playing in a group where at least 50% of subjects have a metacognitive score lower than the sample mean, reduces the chances of choosing C (i.e., cooperate) by almost 0.6% (Model 6). A one-point increase in the score of social metacognition for a subject playing in a group where fewer than 50% of subjects have a metacognitive score lower than the sample mean, reduces the chances of choosing C by 2.2%. Whereas Models 1

and 2 suggest that an increase in the average level of social metacognition in a group increases that chances of choosing C (i.e., it passes from -0.022 to -0.006), Models 6 and 7 show that the distribution of social metacognition within the group is equally important. A decrease in the percentage of people in the group with a social metacognition lower than the sample mean also increases the chances of choosing C (i.e., it passes from -0.022 to -0.006).

From this first set of analyses on the rounds with recommendations, we can summarise the results as follows. First, the effect of recommendations on the likelihood of choosing a C strategy is in line with Duffy and Feltovich (2010). In fact, subjects follow the recommendations regardless of their social metacognition and the level of social metacognition in the group. Second, there is a U-shaped relationship between subjects' metacognitive ability and their propensity to cooperate. We speculate about this result as follows. On one hand, individuals with low SM might know they are not able to understand what opponent subjects may play, so they might blindly choose the C strategy. On the other hand, individuals with high SM might be more likely to play C when they understand it is more profitable–given the opponent predicted choice (recall that the probability of receiving a C recommendation is 2/3). Individuals with SM levels in between neither follow a strategy nor possess the ability to understand others and what to play. Hence, they play C less frequently than the other two types of subjects: (i) people who know they are not able, and (ii) people who correctly know to be highly able to predict the opponent's choice. Third, in groups with a relatively high level of social metacognition, the subjects' metacognitive score has nil, or negative but very low, effect on the chances of playing C. This almost nil result could be a consequence of the already high level of social metacognition in a group. More clearly, an increase in the subjects' social metacognition can bring no results in playing C because the group is already correlating their strategies. This represents our main hypothesis and we confirm it in the next section when we analyse the results in the rounds *without* recommendations.

4.2 Aggregate behaviour without recommendations: the choreographing role of social metacognition

In order to answer the key question of the study, we conduct further analyses at the aggregate level for the rounds where participants do not receive any recommendation. A first regression model studies if social metacognition affects whether participants play a correlated equilibrium, as defined by the four inequalities described in section 3.2 above. The dependent variable is a dichotomous variable of value 1 if the inequalities are satisfied and 0 otherwise.

Table 3 reports the results of two alternative model specifications, differing in which explanatory variables are included. In the first model specification, the independent variables are the average score of social metacognition in the group and the round number. Additionally, we include the proportion of female participants and the proportion of young subjects in the group (i.e., $20 \leq age \leq 30$) together with the group size as control variables. In the second model specification, rather than focusing on the average level of social metacognition in the group, we use as key predictor the fraction of participants in a group whose social metacognition is higher than the sample median. We do not include both measures of social metacognition at the same time to avoid multicollinearity, and we do not use session fixed-effects to avoid overcorrecting for session effect. Standard errors are adjusted for clustering by group.

The table provides evidence that, as rounds proceed further, participants converge towards a correlated equilibrium. In particular, one further round increases the chances of satisfying the inequalities that define a correlated equilibrium by 0.9%. Separate analyses with an interaction term between social metacognition and round number do not provide evidence for an interaction effect. Moreover, a one-unit increase in the average score of social metacognition in the group increases the chances of playing a correlated equilibrium by 0.8%. Although interesting, these results do not immediately support the main hypothesis of the current study as the four inequalities discussed above define a generic correlated equilibrium, therefore, including the Nash equilibria and "bad" correlated equilibria, whose payoff is lower than that of the Nash equilibrium in mixed strategies.

We then conduct a chi-squared test to check if the realised distribution of plays is in line with

Table 3: Results of regression model. Dependent variable: 1 if participants played a correlated equilibrium

	(1)	(2)
Round number	0.009**	0.009**
	(0.004)	(0.004)
Mean group SM	0.008*	
	(0.004)	
Proportion high-SM		0.255
subjects		(0.156)
Proportion female	-0.179	-0.302**
subjects	(0.103)	(0.128)
Proportion young	-0.010	-0.019
subjects	(0.090)	(0.085)
Session size (10	0.133**	0.172**
subects)	(0.055)	(0.065)
Frequency CE	0.415	0.415
Observations	260	260
R-squared	0.034	0.033

SM stands for social metacognition.
CE stands for correlated equilibrium.
Standard errors clustered on session in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the theoretical distribution generated by a mixed-strategy Nash equilibrium where the outcomes are 16% for (D, D), 48% for (C, D) and (D, C), and 36% for (C, C). When we consider the whole sample, the results align to Duffy and Feltovich's findings: the test strongly rejects the null hypothesis that behaviour is generated by i.i.d. mixed-strategy equilibrium play ($p = 0.043$) when (C, D) and (D, C) outcomes are disaggregated, but not when they are pooled ($p = 0.195$). We then conducted the same test separately for groups with a high proportion of metacognitive subjects and groups with a low proportion of metacognitive subjects. The results show that we have enough evidence to rejects the null hypothesis that behaviour in the highly metacognitive groups is generated by i.i.d. mixed-strategy equilibrium play, both when (C, D) and (D, C) outcomes are disaggregated ($p = 0.004$) and when they are pooled ($p = 0.057$). In contrast, for the groups with poor social metacognition we cannot reject the null hypothesis that behaviour is generated by i.i.d. mixed-strategy equilibrium play, both when (C, D) and (D, C) outcomes are disaggregated ($p = 0.178$) and when they are pooled ($p = 0.214$). This result supports the hypothesis that subjects with high social metacognition do not play a Nash equilibrium, and

that subjects with low social metacognition play, instead, according to a mixed strategy Nash equilibrium.

The hypothesis behind the subsequent analyses is that over time subjects with high metacognitive ability - left playing the game without recommendations - can converge their aggregate behaviour towards a correlated equilibrium. More clearly, as our main hypothesis testing, we want to investigate whether social metacognition at the group level predicts the *distance* from the suggested correlated equilibrium distribution. We therefore perform a regression with the distance in absolute value from the given correlated equilibrium as dependent variable. In particular, we record over time the frequency at which participants end up in a given cell of the matrix and observe whether over time there is a convergence in distribution. For instance, consider a group of six participants and suppose that in the first round all the three pairs end up playing C. Hence, after this round the frequency would be 1 for the cell (C, C). Suppose for the second round that the result is one pair in (C, C) and two pairs in (D, D). For this example, at the end of the second round, the frequency distribution would be $2/3$ for (C, C) and $1/3$ for (D, D).

We define with $f_{tg}(C, C)$, $f_{tg}(C, D)$, $f_{tg}(D, C)$, $f_{tg}(D, D)$ the cumulative frequencies for each group g computed in the round t. Moreover, let us consider the correlated equilibrium (suggested in previous rounds) where the distribution is $1/3$ for each of the cells (C, C), (C, D), (D, C), and 0 for the cell (D, D). Our dependent variable $CE_distance_{tg}$, for the round t and group g, is then the distance from the actual distribution coming from the play of the subjects in the group and the theoretical distribution of the correlated equilibrium as above, and it is defined as:

$$CE_distance_{tg} = |f_{tg}(C, C) - 1/3| + |f_{tg}(C, D) - 1/3| + |f_{tg}(D, C) - 1/3| + |f_{tg}(D, D) - 0| \quad (1)$$

Table 4 presents two alternative model specifications, differing in which explanatory variables are included. In the first specification, the independent variables are the average score of social metacognition in the group and the round number. Additionally, we include the proportion of female participants and the proportion of young subjects in the group (i.e., $20 \leq age \leq 30$) together with group size as control variables. The second model specification has as main

predictor the fraction of participants in a group whose social metacognition is higher than the sample median. Separate analyses with an interaction term between social metacognition (with

Table 4: Results of regression model. Dependent variable: Distance from the distribution of the (previously) suggested CE

	(1)	(2)
Round number	-0.006	-0.006
	(0.004)	(0.004)
Mean group SM	-0.020**	
	(0.009)	
Proportion high-SM subjects		-0.009***
		(0.002)
Proportion female subjects	0.266**	0.709***
	(0.120)	(0.148)
Proportion young subjects	-0.157	-0.146*
	(0.109)	(0.081)
Session size (10 subects)	-0.108	-0.242***
	(0.075)	(0.070)
Mean distance	0.449	0.449
Observations	260	260
R-squared	0.371	0.478

SM stands for social metacognition.
Standard errors clustered on session in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

either measure) and round number, do not provide evidence of an interaction effect.

According to our prediction, the coefficients of the two measures of social metacognition are significant and (more importantly) negative. This suggests that high social metacognition in a group leads to a reduction in the distance from the correlated equilibrium used in the first rounds with recommendations. More precisely, from Model 1, it emerges a negative correlation between the average social metacognition in a group and the distance from the suggested correlated equilibrium. Furthermore, from Model 2, an increase by 1 percentage point in the number of highly metacognitive subjects in a group decreases the distance from the suggested correlated equilibrium by about 0.9 percentage points. For example, the addition of one person with high metacognitive abilities in a group with 10 participants (i.e., an increase by 10 percentage points) reduces the distance from the suggested correlated equilibrium by 9%. This suggests that subjects with higher social metacognitive abilities are better able to interrelate their behaviour. Furthermore, play in groups with a high proportion of female participants is further away from

the suggested correlated equilibrium.

5 Discussion

Economics theory postulates that rationally-behaving individuals make decisions in the attempt to maximise their well-being. Consequently, a rational policy maker should consider this in defining optimal policies to maximise social welfare. The heavily quoted Nash equilibrium can aid, as it allows to identify the possible results of social interactions when individuals cannot correlate their strategies. However, the result of a group interaction cannot be detached from the abilities and personal characteristics of the single individuals shaping it. Psychology theories can aid in this regard admitting that individual differences and characteristics could lead to different group compositions and divergent results from social strategic interactions. Also, the group itself could be the result of specific characteristics common to the individuals who are part of it. This paper is a piece of evidence in this regard.

The present study starts from the previous considerations and aims to investigate the link between two theoretical concepts: social metacognition and the correlated equilibrium. Social metacognition is the ability to monitor and control one's own and other's cognitive processes. Creating a representation of the mental states of others, it enables individuals to select responses or actions in social environments. As regulatory process, it helps interacting individuals reach an understanding of each other and regulate accordingly cognitive and metacognitive processes (Volet et al., 2009b). The correlated equilibrium is a solution concept possibly payoff-enhancing with respect to the Nash equilibrium, but with the "inconvenience" of requiring the presence of a *choreographing* mechanism.

We start our conjecture arguing that individuals with high social metacognitive skills could rely on their abilities to interrelate their decisions on better outcomes. We hypothesise that the construct of social metacognition is a good candidate in explaining how individuals can play a correlated equilibrium, even in the absence of a third party. Hence, we postulate that social metacognition works *as if* it is the correlation device.

The main findings combined with prior research provide an insightful preliminary picture

of the role of social metacognition in game theory. In the first set of analyses, we find that social metacognition is a significant predictor of subjects' strategy choices. Subjects with higher metacognitive skills play differently according to the opponents' metacognitive level and group size.

Subsequent analyses show that groups characterised by a higher level of social metacognition more likely converge to the suggested correlated equilibrium in the game, which is payoff-enhancing compared to the Nash equilibrium in mixed strategy. This suggests that individuals with higher social metacognitive abilities are better able to interrelate their behaviour whenever they interact with other socially metacognitive individuals. This result adds to the literature showing the importance of taking into account individual differences in understanding the rationality of behaviour (see e.g., De Neys et al., 2011).

The results are not exhaustive in accounting for individual differences. For instance, the current experiment does not account for the role of cognitive skills. It is possible that low cognitive skills, even if accompanied by high social metacognitive capacities, might not suffice to learn the game and coordinate behaviour with that of the other players. Cognition is an important variable that has been linked to metacognition (Scarampi, 2018) and is associated with behaviour in non-cooperative games (De Neys et al., 2011). Hence, it is likely that high cognitive abilities interact with metacognitive skills, leading to better performance at the game. Future empirical work might look more in depth at cognition as moderating variable of the relationship between metacognition and decision-making in strategic interactions.

Despite the limitations, we believe that our study constitutes an advancement in understanding the determinants of behaviour in strategic interactions and represents a starting point for further developments. It has been stressed that institutions and policies depend upon the use of correlational devices (Gintis, 2009). As a consequence, the society as a whole can benefit from decisions made by agents that rationally follow correlated signals. Since considerably vast policy decisions are based on Nash equilibrium solutions, a more in-depth study on how individual differences and group compositions lead to equilibria different from the Nash equilibrium outcome appear extremely important.

References

Arifovic, J., Boitnott, J. F., and Duffy, J. (2019). Learning correlated equilibria: An evolutionary approach. *Journal of Economic Behavior & Organization*, 157:171 – 190.

Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1(1):67–96.

Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics*, 4(6):1236–1239.

Brown, J. M., Miller, W. R., and Lawendowski, L. A. (1999). The self-regulation questionnaire. In VandeCreek, L. and Jackson, T. L., editors, *Innovations in clinical practice: A source book*, volume 17, page 281–292. Professional Resource Press, Sarasota, FL.

Camerer, C. F., Loewenstein, G., and Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, 43(1):9–64.

Carruthers, P. and Smith, P. (1996). *Theories of Theories of Mind*. Cambridge University Press, Cambridge.

Cason, T. N. and Sharma, T. (2007). Recommended play and correlated equilibria: An experimental study. *Economic Theory*, 33(1):11–27.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2):245–76.

Costello, A. B. and Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7):1–9.

Curry, O. and Chesters, M. J. (2012). 'Putting ourselves in the other fellow's shoes': The role of 'Theory of Mind' in solving coordination problems. *Journal of Cognition and Culture*, 12:147–159.

Davies, M. and Stone, T. (1995a). *Folk Psychology*. Blackwell, Oxford.

Davies, M. and Stone, T. (1995b). *Mental Simulation*. Blackwell, Oxford.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1):113–126.

De Freitas, J., Thomas, K., DeScioli, P., and Pinker, S. (2019). Common knowledge, coordination, and strategic mentalizing in human social life. *Proceedings of the National Academy of Sciences*, 116(28):13751–13758.

De Neys, W., Novitskiy, N., Geeraerts, L., Ramautar, J., and Wagemans, J. (2011). Cognitive control and individual differences in economic ultimatum decision-making. *PLoS One*, 6(11):1–9.

Dimitrijević, A., Hanak, N., Altaras Dimitrijević, A., and Jolić Marjanović, Z. (2018). The mentalization scale (ments): A self-report measure for the assessment of mentalizing capacity. *Journal of Personality Assessment*, 100(3):268–280.

Duffy, J. and Feltovich, N. (2010). Correlated equilibria, good and bad: An experimental study. *International Economic Review*, 51(3):701–721.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, 13(4):277–287.

Efklides, A. (2014). How does metacognition contribute to the regulation of learning? an integrative approach. *Psychological Topics*, 23(1):1–30.

Fehr, D. and Huck, S. (2016). Who knows it is a game? on strategic awareness and cognitive ability. *Experimental Economics*, 19(4):713–726.

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

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y.-W., Warren, F., Howard, S., Ghinai, R., Fearon, P., and Lowyck, B. (2016). Development and validation of a self-report measure of mentalizing: The reflective functioning questionnaire. *PLoS One*, 11(7):1–28.

Foster, D. P. and Vohra, R. V. (1997). Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1):40–55.

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599):2213–2223.

Frith, C. D. and Frith, U. (1999). Interacting minds–a biological basis. *Science*, 286(5445):1692–1695.

Frith, C. D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4):531–534.

Gallagher, H. L., Jack, A. I., Roepstorff, A., and Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage*, 16(3 Pt 1):814–21.

Gintis, H. (2009). *The Bounds of Reason. Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press, Princeton, NJ.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language*, 1(2):158–171.

Hart, S. and Mas-Colell, A. (2000). A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150.

Heal, J. (1986). Replication and functionalism. In Butterfield, J., editor, *Language, Mind and Logic*, pages 135–150. Cambridge University Press, Cambridge.

Jolliffe, D. and Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of Adolescence*, 29(4):589–611.

Jost, J. T., Kruglanski, A. W., and Nelson, T. O. (1998). Social metacognition: an expansionist review. *Personality and Social Psychology Review*, 2(2):137–54.

Lecce, S., Bottiroli, S., Bianco, F., Rosi, A., and Cavallini, E. (2015). Training older adults on theory of mind (tom): Transfer on metamemory. *Archives of Gerontology and Geriatrics*, 60(1):217–226.

McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences of the United States of America*, 98(20):11832–11835.

Myerson, R. B. (1991). *Game Theory: Analysis of Conflict*. Harvard University Press.

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

Powell, J. L., Grossi, D., Corcoran, R., Gobet, F., and Garcia-Finana, M. (2017). The neural correlates of theory of mind and their role during empathy and the game of chess: A functional magnetic resonance imaging study. *Neuroscience*, 355:149–160.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, 22(4):1694–703.

Robalino, N. and Robson, A. (2016). The evolution of strategic sophistication. *American Economic Review*, 106(4):1046–72.

Scarampi, C. (2018). *Age differences in the effects of metacognition on financial decision-making.* (PhD thesis). University of Bath, Bath, UK.

Schraw, G. (2001). Promoting general metacognitive awareness. In Hartman, H., editor, *Metacognition in learning and instruction: Theory, Research and practice*, pages 3–16. Kluwer, Norwell, MA.

Shteynberg, G., Hirsh, J. B., Bentley, R. A., and Garthoff, J. (2020). Shared worlds and shared minds: A theory of collective learning and a psychology of common knowledge. *Psychological Review*, 127(5):918–931.

Volet, S., Summers, M., and Thurman, J. (2009a). High-level co-regulation in collaborative learning: How does it emerge and how is it sustained? *Learning and Instruction*, 19(2):128–143.

Volet, S., Vauras, M., and Pekka, S. (2009b). Self- and social regulation in learning contexts: An integrative perspective. *Educational Psychologist*, 44(4):215–226.

Worthington, R. L. and Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6):806–838.

Appendix A. Supplementary material

Social Metacognition Inventory

A total of 73 items were obtained by adapting items from existing measures related to social metacognition (Brown et al., 1999; Davis, 1983; Dimitrijević et al., 2018; Fonagy et al., 2016; Jolliffe and Farrington, 2006) and inventing some extra statements to cover aspects of the construct that we wanted to measure. SMI items were first piloted with Psychology doctoral students and lecturers who rated the clarity and readability of the items and their appropriateness to measure social metacognition. The instrument was then revised by the first author, who discarded the thirteen items that were judged as most problematic (i.e., poorly phrased, vague, or not clearly measuring the construct of social metacognition).

One of the issues with self-report measures of social metacognition is that to fill in the questionnaire, the respondents have to rely on their mentalising ability, which can be affected by cognitive biases and result in misattributing mental states to others. In the attempt to get around the problem of the "knowledge bias" - i.e., the failure to subtract one's own unique knowledge about the situation when attributing mental states to others (Carruthers and Smith, 1996; Davies and Stone, 1995a,9; Gordon, 1986; Heal, 1986) - a set of items were scored with polar-scoring, whereas other items were scored with a median-scoring method. In the polar-scoring method, the strongest agreement is associated with the highest score of social metacognition (or the lowest score for reverse-coded items). Sample statements are "*I'm often curious about the meaning behind others' actions*" and "*If I'm sure about something, I don't waste time listening to other people's arguments*" respectively. In line with the suggestions made by Fonagy et al. (2016), responses in the median-scoring method should reflect an awareness of the opaqueness of mental states. A sample items is "*I can tell how someone is feeling by looking at their eyes*". Items in this set were thus rescored so that the median score (i.e., 4) corresponded to the highest score of social metacognition and the extreme score to the lowest scores.

A total score of social metacognition was obtained by summing the scores obtained by par-

ticipants at all the items retained from the Exploratory Factor Analysis described below.[4]

SMI validation

In order to study the factor structure of the Social Metacognition Inventory, an Exploratory Factor Analysis (EFA) was performed with SPSS version 22. First of all, we computed the Pearson correlation matrix between the 63 items of the SMI. A few items with only one or two correlations with other items exceeding .3 were identified and removed. The anti-image correlation matrix suggested reasonable factorability. Furthermore, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy on the remaining items of the SMI was .69 and the Bartlett's test of sphericity was significant ($\chi^2(1225) = 2998.13$, $p < .001$), indicating that the matrix was suited to factor analysis (Worthington and Whittaker, 2006).

A principal axis analysis was performed on the data, following a standard approach to conducting an EFA (Costello and Osborne, 2005; Worthington and Whittaker, 2006). A three-factor and a five-factor competing solutions were initially suggested by Cattell (1966)'s scree test. An Oblimin (oblique) rotation was performed to clarify the data structure. An oblique rotational method was chosen as the obtained factors were hypothesised to be related, based on previous theoretical considerations (Efklides, 2008). The EFA was also used to inform the exclusion and retention of items (Worthington and Whittaker, 2006). A priori criteria for factor loadings were set to 0.32. Items were discarded if they either did not load well onto any factor or had significant cross-loadings onto other factors.

The three-factor model was chosen as final solution after considering the internal consistency of the obtained factors and the interpretability of the obtained factor solutions (a more detailed description of item extraction and differences between the two models can be found in the Supplementary Material; Worthington and Whittaker, 2006). The pattern matrix obtained for the three-factor solution consisted of 43 items. It revealed a first factor consisting of 13 items, a second factor consisting of 14 items, and a third factor consisting of 16 items. The factor

[4] A larger sample of participants was used with the aim of increasing the power of the Exploratory Factor Analysis and the validation of the Social Metacognition Inventory. Overall, data were collected from 122 individuals (age range 20-79, $M = 43.09$ years, $SD = 19.80$; 69 female).

loadings of the items of the SMI are displayed in Table 1.

Table A.1: Pattern Matrix for the Social Metacognition Inventory

Item	Factor 1	Factor 2	Factor 3
41. I often remind others to contribute their ideas.	0.67		
73. I offer to help others during a group work.	0.67		
26. I call in others for help when I need it.	0.64		
27. I ask for clarification if I do not understand something.	0.63		
35. When working on a group project, I often help others who have difficulties in understanding the group task.	0.61		
25. When working in a group, I often give feedback to contributions made by others.	0.59		
3. When working on a collaborative project I discuss thoughts with other group members.	0.57		
38. When making plans with other people, I try to make sure our plans are realistic.	0.56		
19. When working on a collaborative project I compare thoughts with other group members.	0.52		
56. When working with other people, I try to make sure we set learning goals and allocate time for various activities.	0.51		
72. When working on a group project, I often try to work with others to complete our task.	0.45		
40. When working on a group project, I often feel pleased if others remind me of the time remaining to finish our work.	0.45		
5. Before starting working on a group project, we set goals to guide what steps we would take.	0.42		
39. When working on a group project, I often try to remind others of the time remaining to finish our work.	0.41		

Item	Factor 1	Factor 2	Factor 3
22. When working on a collaborative project, I try to make sure we all make efforts to achieve our set goals.	0.38		
62. I am comfortable working with a group.	0.38		
59. It takes me a long time to understand other people's thoughts and feelings.		0.68	
52. I can often understand how people are feeling even before they tell me.		0.61	
33. I have trouble figuring out my friends' feelings.		0.59	
16. My gut feeling about what someone else is thinking is usually very accurate.		0.57	
53. People's thoughts are a mystery to me.		0.56	
32. Other people's thoughts and feelings are confusing to me.		0.55	
9. I can tell how someone is feeling by looking at their eyes.		0.51	
15. I can make good predictions of other people's behaviour when I know their beliefs and feelings.		0.50	
1. I can easily deduce someone's intentions.		0.50	
12. It's really hard for me to figure out what goes on in other people's heads.		0.47	
58. Understanding what's on someone else's mind is never difficult for me.		0.46	
20. I usually know exactly what other people are thinking.		0.35	
31. I can mostly predict what someone else will do.		0.33	
65. I often think about other people and their behaviour.			0.68
64. I find it important to understand reasons for behaviour.			0.63
49. I do not like to waste time trying to understand in detail other people's behaviour.			0.60

Item	Factor 1	Factor 2	Factor 3
57. Understanding the reasons for people's actions helps me to forgive them.			0.55
2. When interacting with someone else I try to understand their thoughts.			0.55
63. When I'm upset at someone, I usually try to "put myself in his shoes" for a while.			0.53
11. I'm often curious about the meaning behind others' actions.			0.50
17. In an argument, I keep the other person's point of view in mind.			0.43
60. Two people can see the same image and interpret it differently.			0.42
43. I try to look at everybody's side of a disagreement before I make a decision.			0.42
54. I pay attention to the impact of my actions on others' feelings.			0.41
18. I believe that people can see a situation very differently based on their own beliefs and experiences.			0.39
37. If I'm sure I'm right about something, I don't waste time listening to other people's arguments.			0.33
28. I take into account the ideas and suggestions of others.			0.32
Eigenvalue	7.98	4.30	3.38
Percentage of variance	18.57	10.00	7.86
Cronbach's Alpha	0.87	0.83	0.83

Inspection of the item contents revealed that the three factors were related to the expected dimension and represented the subscales hypothesised based on previous theoretical considerations (Efklides, 2008). The first factor was robust, with a high eigenvalue of 7.98, and it accounted for 18.57% of the variance in the data. From Table 1 it emerges that these items all relate to the ability to monitor and control cognitive performance in the context of social interactions. This factor loaded onto reported levels of ability to adapt actions and behaviour to different social

environments and co-regulate performance in collaborative contexts. As a consequence, in light of the contents of the items, the factor was labelled "Metacognitive Skills". The second factor had an eigenvalue of 4.30 and accounted for 10% of the variance in the data. Items loading onto this factor related to the ability to take into account and make an estimate of the mental states of others and think about other people's actions and outcomes. This factor was thus labelled "Metacognitive Judgements". The third factor had an eigenvalue of 3.38 and accounted for 7.86% of the variance in the data. This factor related to the ability to build knowledge on social interactions, enhancing in turn the ability to understand how to behave in the specific environment. In line with the reference model of social metacognition proposed by Efklides (2008), this factor was labelled "Metacognitive Knowledge".

Internal consistency for each scale was examined using Cronbach's alpha. The three factors demonstrated a high degree of internal consistency. The α coefficient for metacognitive skills was 0.87, while the coefficients for metacognitive judgements and metacognitive knowledge were both 0.83. The Cronbach's α for the entire questionnaire was 0.78. Overall, the factor analysis indicated three distinct factors underlying participants' responses at the SMI, all internally consistent.