Disclosure of Belief-Dependent Preferences in a Trust Game

Giuseppe Attanasi (Sapienza University of Rome)

Pierpaolo Battigalli (Bocconi University and IGIER, Milan)

Elena Manzoni (University of Bergamo)

Rosemarie Nagel (ICREA, Universitat Pompeu Fabra, Barcelona GSE)

Online Appendix A

In this appendix we provide:

    A.1   The translation from Italian of the experimental instructions. Sentences written in Italics within brackets are not part of the Instructions; they are just clarifying comments for the reader.

    A.2   Additional comments to the experimental design.

A.1 INSTRUCTIONS

Welcome to the Experimental Laboratory of Bocconi University, Milan. Thank you for participating in this experiment.

What you will earn at the end of the experiment depends on your decisions and those of the other participants.

If you have any questions, you can ask aloud now. We will answer your question publicly.

From this point on, you are no longer allowed to ask questions publicly. If you have a question after the experiment has started, please raise your hand and we will answer your question privately.

This experiment consists of **three** *[two for NoQ]* phases. For each phase, you will receive specific instructions before its beginning.

PHASE 1

Now we explain in detail the instructions of Phase 1.

You will be randomly paired with another participant.

In each pair, one participant will be in role **A** and the other will be in role **B**.

You will never know the identity of the participant being paired with you, nor he/she will ever know who you are.

There are 20 participants, thus there will be 10 participants in role **A** and 10 in role **B**.

You will be told which is your role before Phase 1 will start.

Now we explain what you are asked to do in Phase 1.

CHOICES

In each pair,
- the participant in role A chooses between two actions: **DISSOLVE** or **CONTINUE**;
- the participant in role B chooses between two actions: *SHARE* or *TAKE*.

The two participants choose simultaneously. Hence B, when choosing, does not know if A has chosen **DISSOLVE** or **CONTINUE** and vice versa.

Each participant will be paid according to his/her choice and to the choice of the participant paired with him/her, as described in the following Table of Payments.

	A receives	B receives
A **DISSOLVE**, (B *SHARE* or *TAKE)*	1€	1€
A **CONTINUE**, B *SHARE*	2€	2€
A **CONTINUE**, B *TAKE*	0€	4€

Table of Payments

Explanation of the Table of Payments:

At first, we assign 2€ to the pair.
If A chooses **DISSOLVE**, the 2€ are shared equally between the pair (1€ to A, 1€ to B), regardless of whether B has chosen *SHARE* or *TAKE*.
If A chooses **CONTINUE**, we add other 2€ to the pair. At this point, there are 4€ to be divided. In this case, the division depends on B's choice: if B has chosen *SHARE*, then A receives 2€ and B receives 2€; if B has chosen *TAKE*, then A receives 0€ and B receives 4€. Recall that the game is simultaneous: B chooses *SHARE* or *TAKE* without knowing what A has chosen and vice versa.

ESTIMATES and CONJECTURES

Before you choose the action, we will ask you to make an estimate (if you have role A) or two conjectures (if you have role B). In particular,

- A's <u>ESTIMATE</u>:
We will ask each A to guess the percentage of participants B choosing *SHARE*.
Recall that there are 10 participants having the role B, hence the only possible percentages are 0% (= 0/10 choosing *SHARE*), 10% (= 1/10 choosing *SHARE*), 20% (= 2/20), 30% (= 3/10), 40% (= 4/10), 50% (= 5/10), 60% (= 6/10), 70% (= 7/10), 80% (= 8/10), 90% (= 9/10), 100% (= 10/10 choosing *SHARE*). Every participant A has to indicate one of these percentages.
The accuracy of A's estimate depends on the choice made by all participants B.

- B's <u>CONJECTURE 1</u>:
We will ask each B to guess the estimate given by the participant A paired with him/her about the percentage of participants B choosing *SHARE*. This means that B has to guess whether the participant A paired with him/her said that 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 100% of participants B choose *SHARE*.
The accuracy of B's "conjecture 1" depends only on the estimate made by the participant A paired with him/her.

- B's <u>CONJECTURE 2</u>:
We will ask each B to guess the action chosen by the participant A paired with him/her (**DISSOLVE** or **CONTINUE**).
The accuracy of B's "conjecture 2" depends only on the choice made by the participant A paired with him/her.

Information after Phase 1

At the end of Phase 1, you will not receive any information about choices, estimates and conjectures of anybody.
At the end of the whole experiment, you will be told the choice made by the participant who is paired with you and whether your estimate (if you have the role A) or your conjectures (if you have the role B) were correct.

Payments for Phase 1

<u>Choices</u>: at the end of the whole experiment, you will be paid (according to the Table of Payments above) based on your choice and on the choice made by the participant who is paired with you.
<u>Estimate</u>: (only for role A): at the end of the whole experiment, if your estimate was correct, you will receive €5, otherwise you will receive € 0.
<u>Conjectures</u>: (only for role B): at the end of the whole experiment, if both your conjecture 1 and conjecture 2 were correct, you will receive €5, otherwise (in every other possible case) you will receive €0.

Therefore, in this phase you are asked to perform two tasks:
- a choice and an estimate (if your role is A);
- a choice and two conjectures (if your role is B).

Now the computer will randomly select your role (A or B).
Then, Phase 1 will start.

PHASE 2

[Only for QnoD and QD]

Your role is the same as in Phase 1.

The following refers only to participants in role B.

Participants in role A have in front of them the same instructions, in order to be informed about what we are asking to participants in role B.

Consider this hypothetical situation:

- Suppose that you have been paired with another participant A to play the same game described in the Table of Payments of Phase 1.

- Regardless of what you have actually chosen in Phase 1, suppose that new A chose **CONTINUE** and you chose *TAKE*, hence you got €4 and left A with €0 in his/her pocket.

- Suppose that, after obtaining €4, you have the opportunity to give back part of this €4 to A. What you give back to him/her can differ according to his/her guess about you choosing *SHARE*.

Look at the *Hypothetical Payback Scheme* in Sheet 1. In the left column, there are 11 possible A's guesses about you choosing *SHARE*.

Keeping in mind that we are assuming that the A paired with you has chosen **CONTINUE** and that you have chosen *TAKE*, we ask you to indicate how much (if any) of the €4 obtained you are willing to give him/her back. You have to indicate it for each possible guess of A about you choosing *SHARE*.

Please, fill in accordingly the right column of the *Hypothetical Payback Scheme* in "Role B - Sheet 1" and copy the entered values in the one that will appear on your computer screen once Phase 2 will start. Then answer the qualitative questions in "Role B - Sheet 2".

<u>Important</u>:

- You will not receive any payment for the values entered in the *Hypothetical Payback Scheme*, nor these values will bring any monetary transfer to any of the participants in role A;

[In QnoD]
- Once completed, your *Hypothetical Payback Scheme* will not be forwarded to anyone.

[In QD]
- Once completed, your *Hypothetical Payback Scheme* will be forwarded to a randomly chosen participant in role A.

Role B - Sheet 1

[Only for QnoD and QD]

Suppose A chose CONTINUE and you have chosen TAKE, hence getting €4, but now you can give part of this amount back to A. How much would you give, if A expected you to SHARE with some probability?

Before choosing **CONTINUE**, A thought that you would have chosen *SHARE* with probability:	You would give back to A the sum of:
0%	€
10%	€
20%	€
30%	€
40%	€
50%	€
60%	€
70%	€
80%	€
90%	€
100%	€

Hypothetical Payback Scheme

In each line of the right column of the Hypothetical Payback Scheme above (for each guess of A about you choosing *SHARE*), you have to enter a value between €0.00 and €4.00. Each value can be entered in decimal form, up to two decimal digits.

Role B - Sheet 2

[Only for QnoD and QD]

We ask you to comment on the values you entered in the Hypothetical Payback Scheme.

The following comments will not be passed on to any participant in the experiment.

- Explain the meaning of the values you entered in the Hypothetical Payback Scheme.

_____

_____

_____

_____

_____

_____

_____

_____

- Did you enter these values according to a specific feeling?   Yes ☐   No ☐

Keep on answering only if the answer to the previous question is "Yes".

- Is this feeling dependent on your partner's guess about you choosing *SHARE*?

Yes ☐   No ☐

Keep on answering only if the answer to the previous question is "Yes".

- What kind of relationship is there between this feeling and your partner's guess about you choosing *SHARE*?

_____

_____

_____

_____

_____

_____

_____

Role A - Sheet 1

[Only for QD, after all participants in role B have filled in their Hypothetical Payback Scheme: At this point participants in role A receive on their computer screen a filled-in Hypothetical Payback Scheme by a randomly chosen participant B (different for each participant A)]

The following refers only to participants in role A.

Participants in role B have in front of them the same instructions, in order to be informed about what we are asking to participants in role A.

Now that all participants in role B have filled in their Hypothetical Payback Scheme, you will see on the screen of your computer the one filled in by a randomly chosen participant B (different for each participant A).

We kindly ask you to copy in the Scheme below the values entered by the randomly chosen participant B and to answer the questions in "Role A - Sheet 2".

In case you choose **CONTINUE**, your guess that B would choose *SHARE* could be:	The sum (between €0.00 and €4.00) that B would give you back in case you would choose **CONTINUE** and he/she would choose *TAKE* is:
0%	€
10%	€
20%	€
30%	€
40%	€
50%	€
60%	€
70%	€
80%	€
90%	€
100%	€

Hypothetical Payback Scheme filled in by a randomly chosen participant B

Role A - Sheet 2

[Only for QD, after all participants in role B have filled in their Hypothetical Payback Scheme: At this point participants in role A receive on their computer screen a filled-in Hypothetical Payback Scheme by a randomly chosen participant B (different for each participant A)]

We ask you to comment on the values entered in the Hypothetical Payback Scheme by the randomly chosen participant B.

The following comments will not be passed to any participant in the experiment.

- Write what you think is the meaning of the values that the randomly chosen participant B has entered in the Hypothetical Payback Scheme that you have been forwarded.

_____

_____

_____

_____

_____

_____

- Do you think that the randomly chosen participant B entered these values according to a specific feeling?

Yes ☐   No ☐

Keep on answering only if the answer to the previous question is "Yes".

- According to the values he/she has entered in the Hypothetical Payback Scheme, if the randomly chosen participant B were asked to play the game in Phase 1 with you, what kind of feeling do you think he/she would feel had you chosen **CONTINUE** and had he/she chosen *TAKE*?

_____

_____

_____

_____

- Do you think that his/her feeling could depend on your guess about he/she choosing *SHARE*? If yes, in which way?

_____

_____

_____

_____

_____

PHASE 3 *[PHASE 2 for NoQ]*

[Participants have instructions of Phase 1 in front of them.]

[Only for QnoD and QD: Participants in role B also have their Sheet 1 of Phase 2 with the filled-in Hypothetical Payback Scheme in front of them.]

[Only for QD: Participants in role A also have their Sheet 1 of Phase 2 with the received filled-in Hypothetical Payback Scheme in front of them.]

Your role is the same as in Phase 1 and Phase 2 *[as in Phase 1 for NoQ]*.

Now we explain in detail what you will be asked to do in Phase 3 *[Phase 2 for NoQ]*.

The participant paired with you during Phase 3 *[Phase 2 for NoQ]* is not the same as in Phase 1. At the end of Phase 2 *[Phase 1 for NoQ]*, the pairs have been randomly re-matched and you have been paired with another participant, with whom you will interact in Phase 3 *[Phase 2 for NoQ]*.

[Only for QD]

Indeed, during Phase 3, each participant A is paired with the randomly chosen participant B in Phase 2, namely the one from whom A has been forwarded the filled-in Hypothetical Payback Scheme. Each participant A can keep the paired B's filled-in Hypothetical Payback Scheme in paper form ("Role A - Sheet 1" of Phase 2) in front of him/her during Phase 3.

[Only for QnoD and QD]

Each participant B can keep his/her previously filled-in Hypothetical Payback Scheme in paper form ("Role B - Sheet 1" of Phase 2) in front of him/her during Phase 3. Note that the payments in Phase 3 do not depend on the values that a participant B has entered in the Hypothetical Payback Scheme during Phase 2.

In Phase 3 *[Phase 2 for NoQ]*, you will play the same game played in Phase 1, the one described in the Table of Payments.
In particular, what we will ask you to do is, as in Phase 1,
- a choice and an estimate (if your role is A);
- a choice and two conjectures (if your role is B).

Information after Phase 3 *[Phase 2 for NoQ]*
At the end of Phase 3 *[Phase 2 for NoQ]*, you will not receive any information about choices, estimates and conjectures of anybody in this phase.
At the end of the whole experiment, you will be told the choice made by the participant who is paired with you in Phase 3 *[Phase 2 for NoQ]* and whether your estimate (if you have the role A) or your conjectures (if you have the role B) in Phase 3 *[Phase 2 for NoQ]* were correct.

Payments for Phase 3 *[Phase 2 for NoQ]*
Choices, Estimates and Conjectures will be paid according to the same rules established for Phase 1.

Final Questionnaire - Role B

So far you have played the game described in the Table of Payments twice (once in Phase 1, once in Phase 3 *[Phase 2 for NoQ]*). In both cases, you do not know what the participants you were paired with have chosen.

[For NoQ: see the instructions of Phase 2 in QnoD]

[For QnoD and QD: see the following instructions]

Now we ask you to imagine again the same hypothetical situation described in Phase 2. Please fill in the *Final Hypothetical Payback Scheme*, which has the same structure of the Hypothetical Payback Scheme that you have filled in Phase 2.

<u>**Important**</u>:

- For any line in the *Final Hypothetical Payback Scheme*, you can enter, if you wish, a value which is different from the one entered in the corresponding line of the *Hypothetical Payback Scheme* you have filled in Phase 2;

- You will not receive any payment for the values entered in the *Final Hypothetical Payback Scheme* nor these values will bring any monetary transfer to any of the participants with role A;

[In QnoD]
- As for the Hypothetical Payback Scheme of Phase 2, once completed, your *Final Hypothetical Payback Scheme* will not be forwarded to anyone.

[In QD]
- Differently from the Hypothetical Payback Scheme of Phase 2, once completed, your *Final Hypothetical Payback Scheme* will not be forwarded to anyone.

Suppose A chose CONTINUE and you have chosen TAKE, hence getting €4, but now you can give part of this amount back to A. How much would you give, if A expected you to SHARE with some probability?

Before choosing **CONTINUE**, A thought that you would have chosen *SHARE* with probability:	You would give back to A the sum of:
0%	€
10%	€
20%	€
30%	€
40%	€
50%	€
60%	€
70%	€
80%	€
90%	€
100%	€

Final Hypothetical Payback Scheme

Explain the meaning of the values you entered in the *Final Hypothetical Payback Scheme*.

_____
_____
_____
_____

- Are these values different from those you entered in the Hypothetical Payback Scheme in Phase 2?

Yes ☐   No ☐

Keep on answering only if the answer to the previous question is "Yes".

- Which of the following statements better explains why the values you entered in the *Final Hypothetical Payback Scheme* are different from those you entered in the *Hypothetical Payback Scheme* in Phase 2? [You can indicate more than one motivation.]

I.   At the moment I have *the same* kind of feeling that I expressed by filling in the Hypothetical Payback Scheme in Phase 2. But now that feeling is "*stronger*".   ☐
Why? _____
_____
_____

II.   At the moment I have *the same* kind of feeling that I expressed by filling in the Hypothetical Payback Scheme in Phase 2. But now that feeling is "*weaker*".   ☐
Why? _____
_____
_____

III.   At the moment I *do not* have *the same* kind of feeling that I expressed by filling in the Hypothetical Payback Scheme in Phase 2.   ☐
Which is the new feeling?_____
_____
_____

IV.   [*Only for QD*] Contrarily to the Hypothetical Payback Scheme in Phase 2, this time the Final Hypothetical Payback Scheme *will not* be forwarded to any participant in this experiment.   ☐
Additional Comments: _____
_____
_____

V.   Other motivations   ☐
_____
_____
_____

A.2 COMMENTS ON THE EXPERIMENTAL DESIGN

In this appendix, we comment on some important features of the experimental design and provide motivations for specific design choices.

Relevance of Phase 1. There are two reasons for the initial one-shot interaction in phase 1. First, we want to know how subjects form their beliefs and make their choices without public information about B's answers to the questionnaire in phase 2. This allows us to test for within-subject effects of questionnaire disclosure in QD. Second, we want to let subjects understand the Trust Minigame and the belief-elicitation procedure before B-subjects fill in the questionnaire in phase 2 of $QnoD$ and QD. Indeed, each of the 11 possible guesses for the frequency of *Share* in phase 1 corresponds to one of the 11 rows of the questionnaire of Table 3 of the paper (i.e., the hypothetical payback scheme above), which makes it more salient.

For *NoQ*, phase 1 has been mainly introduced to maintain the same structure as in *QnoD* and *QD*, thereby balancing observations across treatments.

Beliefs Elicitation in Phase 1 and Phase 3. We made several specific design choices about the belief-elicitation procedure, building on previous experimental literature. You can refer to Section 4 of Schotter & Trevino (2014), for a survey on first- and second-order beliefs elicitation in two-player games with belief-dependent motivations.

Charness & Dufwenberg (2006) use the strategy method to elicit the contingent choice of B-subjects in the standard, sequential version of the Trust Minigame. In this respect, our approach is similar; we make subjects play the strategic form of the Trust Minigame (see also Guerra & Zizzo 2004 and Bacharach *et al.* 2007). Due to possible framing effects, there is a subtle difference between (i) presenting subjects with a sequential game and then use the strategy method, and (ii) presenting them---as we do---with a simultaneous game corresponding to the strategic form of the sequential one (cf. Siniscalchi 2016). But we think that our description of the game in the instructions avoids such framing effects.

Differently from Charness & Dufwenberg (2006), and similarly to Guerra & Zizzo (2004) and Bacharach *et al.* (2007), we elicit beliefs before choices. The experimental results by Guerra & Zizzo (2004) suggest that eliciting beliefs first should not change behavior in the subsequent Trust Minigame.

First-order beliefs of A-subjects are elicited as in Charness & Dufwenberg (2006) and follow-up papers on the Trust Minigame (see, e.g., Bracht & Regner 2013). Like them, we ask A to guess how many of the 10 B-subjects in her session would choose *Share*. Since subjects know they are paired randomly, this is a reasonable measure of first-order beliefs.

As for B-subjects, we elicit B's *unconditional* second-order belief of *Share*, while Charness & Dufwenberg (2006) elicit B's second-order belief of *Share* conditional on A choosing *Continue*. The main reason why we elicit unconditional rather than conditional beliefs relates to the questionnaire in phase 2, which has a central role in our design. As explained above, we want to match the probability grid with the 11 possible answers in A's belief-elicitation task. Thus, in order to have a manageable number of rows in the questionnaire, we only have 10 A-B pairs in each session. This is too small a number for making a reliable inference about A-subjects' first-order belief of *Share*, if one considers only those choosing *Continue*.

Our choice is also supported by theoretical considerations. On the one hand, the choice of B in the simultaneous game is equivalent to a contingent plan in the sequential version of the game, and, therefore, it should correlate with his conditional belief. On the other hand, unconditional beliefs are relevant as well, because they reflect how players reason strategically before playing the game.

The theoretical analysis of the paper provides some testable predictions about unconditional beliefs in phases 1 and 3 of *NoQ* and *QnoD* and in phase 1 of *QD*. This also motivates our elicitation of B's unconditional *first*-order beliefs, unlike most previous experimental studies on the Trust Minigame.[1]
For the sake of simplicity, we just elicit a coarse feature of the first-order beliefs of B-subjects, that

---

[1] For an exception, see Regner & Harth (2014). Chang *et al.* (2011) also elicit B's first-order beliefs, although they do not use them in the analysis.

is, the action of the co-player *A* that they deem more likely. For the *B*-subjects who guess *Continue*, the unconditional second-order belief is also a rough estimate of the conditional one (see footnote 36 in the paper). Notice that the payment scheme of *B*'s second-order beliefs requires *B* to guess correctly both the choice and the first-order belief of *A*, which is consistent with the theoretical definition of second-order belief as a joint distribution about the first-order belief and the action of the co-player.

Questionnaire in Phase 2. In both treatments *QnoD* and *QD*, *A*-subjects read and listen to the instructions of phase 2. This is made so that *A*-subjects know the task of *B*-subjects in phase 2, and in *QD* also to help them interpret the disclosed filled-in questionnaire.

The reason for asking *B*-subjects to fill in the questionnaire twice---first on a sheet of paper and then on the computer screen---is to make them think more carefully about their answers. A similar consideration applies to *A*-subjects in *QD*: They see *B*'s answers on their computer screen and they are asked to copy them on a sheet of paper in order to increase their attention.

Finally, we comment on withholding the identity of the recipient of *B*'s filled-in questionnaire in the main treatment, *QD*. In phase 2, we tell subjects as little as possible about phase 3. Although subjects know that there is a phase 3, they do not know how the experiment will continue, hence they do not know if and how their answers to the questionnaire will be used later. Specifically, in phase 2 it is public information that the filled-in questionnaire *will be disclosed* to a randomly-drawn *A*-subject at the end of this phase. But only at the beginning of phase 3 it is made public information within each pair that the randomly-drawn player *A* corresponds to the matched *A*-subject of phase 3. With this, *B*-subjects should not have any obvious incentive to manipulate the beliefs of the recipient of their filled-in questionnaire. On such unexpected data use, see Charness *et al.* (2022), and, in particular, Section 5.

Final questionnaire. When *B*-subjects fill in the final questionnaire, they know that there is no further decision task to execute; therefore, they should not have any incentive to lie. The final questionnaire provides information about B-subjects who did not fill in a questionnaire in phase 2 (in *NoQ*), and allows us to check whether the *B*-subjects who filled in the questionnaire in phase 2 change or confirm their answers (in *QnoD* and *QD*). In the latter case, we cannot reject the hypothesis that subjects truthfully revealed their belief-dependent preferences in phase 2.

Note that, in *QnoD* and *QD*, at the end of phase 3, the experimenter withdraws the phase 2 filled-in questionnaire in paper form, so as to prevent *B*-subjects from looking at their answers of phase 2 when filling in the final questionnaire. Leaving this paper with them could have biased the answers to the final questionnaire.

Additional references

Schotter, A., and I. Trevino (2014): "Belief Elicitation in the Lab," *Annual Review of Economics*, 6, 103-128.
Siniscalchi, M. (2016): "Structural Rationality in Dynamic Games," Econometrica, 90, 2437-2469.