

Problems with Measurement of Trust and Trustworthiness. What Best Predicts Trust Game Outcomes?¹

Anna O. Kuźmińska*

The study (N = 156) explored the predictive power of attitudinal measures of trust and trustworthiness commonly used in surveys, as well as the possibility of replacing them with other instruments. Secondly, it aimed at replicating previous findings (e.g. Gleaser et al., 2000) that cues to other person's group membership affect trusting and trustworthy behaviour. Subjects filled in a number of questionnaires and took part in the Trust Game in which information about partner's nationality was manipulated. Analyses demonstrated the lack of correspondence between attitudinal and past behaviour measures of trust and trustworthiness and behaviour in the Trust Game. Trustworthiness was predicted by social desirability measure. No effects of other person's out-group membership were found in the current study. Cues to identity made participants more trustful and trustworthy towards the other person in comparison to the control condition. The interpretation of the obtained results, as well as the limitations of the study are discussed.

Keywords: trust, trustworthiness, social desirability, cues to identity.

Submitted: 12.11.2016 | Accepted: 28.12.2016

Zaufanie i wiarygodność: pomiar i wpływ informacji o przynależności grupowej

Badanie (N = 156) miało na celu sprawdzenie wartości predykcyjnej kwestionariuszowych miar zaufania w odniesieniu do zachowania w grze oraz możliwości zastąpienia ich przez inne narzędzia. Ponadto celem była replikacja uzyskanych przez innych badaczy wyników, demonstrujących, że informacja o przynależności grupowej drugiej osoby wpływa na ufne i godne zaufania zachowanie w grze. Po wypełnieniu kwestionariuszy, uczestnicy wzięli udział w Grze Zaufania, w której manipulowano informacją na temat przynależności grupowej drugiej osoby. Analiza wyników potwierdziła przewidywania na temat braku związku pomiędzy deklaracyjnymi miarami zaufania i zachowaniem w grze. Wiarygodność przewidywała natomiast miara aprobaty społecznej. Wyniki nie potwierdziły negatywnego wpływu informacji o przynależności do grupy obcej na zaufanie i wiarygodność. Obecność informacji o drugiej osobie spowodowała wyższą ufność i wiarygodność w porównaniu do grupy kontrolnej. Dyskusja odnosi się do potencjalnych przyczyn otrzymanych wyników oraz ograniczeń prezentowanego badania.

Słowa kluczowe: zaufanie, wiarygodność, aprobaty społeczna, przynależność grupowa.

Nadesłany: 12.11.2016 | Zaakceptowany do druku: 28.12.2016

JEL: C18

* **Anna O. Kuźmińska** –PhD, Faculty of Management, University of Warsaw.
Mailing address: Faculty of Management, University of Warsaw, ul. Szturmowa 1/3, 02-678 Warsaw, Poland; e-mail: akuzminska@wz.uw.edu.pl.



1. Introduction

Trust is viewed as a prerequisite of functional social interactions and effective market exchanges (La Porta et al., 1997). Some researchers conceptualize it as an encapsulated interest, where one decides to trust only on the basis of a reasonable suspicion that their partner would benefit from being trustworthy (Berg et al., 1995; Hardin, 2006). Others (Baier, 1986; Rousseau et al., 1998; Leslie, 2004) prefer to consider trust as based on an emotional belief in someone's benevolence, involving an acceptance of vulnerability and capability of betrayal, at the same time distinguishing it from *reliance*, which cannot be betrayed, but merely disappointed. Similarly, Yamagishi & Yamagishi (1994), point out that it is important to differentiate *trust*, which, in their opinion, refers to: "An expectation of partner's goodwill and benign intent" from *confidence*, which is: "An expectation of partner's competence" (p. 131), as well as *assurance* that may be defined as: "Knowledge of the incentive structure surrounding the relationship" (p. 132) and thus approaches Hardin's idea of encapsulated interest.

A related concept and one of similarly unclear definitions, is trustworthiness. Hardin (2006) argues that, although many scientists speak about the decline of trust in Western societies (Canada, United States, Sweden, or United Kingdom) they, in reality, mean that perceived trustworthiness is lower than it used to be. Hence, trustworthiness may be viewed as the extent to which someone is worthy of being trusted. Trust may originate from the perceived trustworthiness of the other party – a belief in their good intentions (Caldwell & Clapham, 2003). Trustworthiness is assessed on the basis of previous experiences, as well as directly and indirectly obtained information about a trustee (Sztompka, 2007).

This unsettled dispute as to what is the exact definition of trust, trustworthiness, and, for that matter, trustfulness, reliance, or confidence, may impact the matter of their measurement. For instance, Yamagishi & Yamagishi (1994) argue that, although Japanese society is often viewed and described as highly trustful, the level of trust there is in fact lower than among Americans. This confusion, in their view,

stems from the fact that what is being measured is not trust, but assurance, which is based on commitment and stable relationships. In any case, lack of trust in the society increases the cost of social exchange and thus, even though trusting may be costly when one deals with an untrustworthy partner, it is generally beneficial for all parties involved (Axelrod, 1984). Research indicated that trust in strangers allows people to get out of the committed relationships, which sometimes incur high opportunity costs, and explore more advantageous alternatives (Hayashi & Yamagishi, 1998). What is more, a trusting attitude is positively related to income, while this relation is negative for trustworthiness (Slemrod & Katuscak, 2005). A similar result was reported by Knack and Keefer (1997), who demonstrated that trust was connected with superior economic performance.

The first aim of this research is, therefore, to take a closer look at the measurement of trust and trustworthiness. In particular, it is investigated whether traditionally used questionnaires predict trustful and trustworthy behaviour in the trust game. The second aim is the replication of the results obtained by other researchers (e.g. Fershtman & Gneezy, 2001; Foddy, Platow, & Yamagishi, 2009; Gleaser et al., 2000), showing that information about other person's group membership affects people's trusting and trustworthy behaviour towards that person, with less trust and trustworthiness expressed towards the out-group members.

2. Problems with Different Types of Trust Measurement

Currently, there are basically two ways in which trust and trustworthiness are measured in research. One method employs surveys and questionnaires, the other game experiments (Kohn, 2008; Hardin, 2006). Both of them have advantages as well as limitations.

2.1. Attitudinal Measures of Trust

Previous research by Gleaser et al. (2000) showed that responses on none of the trust questionnaires used managed to predict trusting behaviour in the Trust Game. Trustfulness in the games, in the Gleaser et al.'s study, was predicted only by an index of questions relating to subjects'

previous trusting behaviour, e.g. “How often do you lend money to your friends?”, while trustworthy behaviour was – surprisingly – predicted by questions about general trust. These results were partly supported by Holm and Danielson (2005), who found, however, that neither attitudinal nor past behaviour measures predicted trust or trustworthiness in games.

Absence of association between attitudes expressed in questionnaires and actual behaviour is not a new discovery. It was first documented by LaPiere (1928), who found that hotel and restaurant owners admitted Chinese couples to their premises, even though most of them had previously stated that “members of the Chinese race” would not be welcome. This lack of correspondence between attitudes and behaviour was subsequently explored in a large number of studies. In a review of the literature, Ajzen and Fishbein (1977) observed that while some studies supported LaPiere’s findings, others found attitudes and behaviour to be correlated. The authors argue that in order for the attitudinal measures to be correct in predicting behaviour, they have to be precise with respect to action, target, and context. This exactly may be the problem with attitudinal measures of trust and trustworthiness. The generality of questions asked, when it comes to target (e.g. “Most people can be trusted”), as well as context (e.g. “I am trustful”), might lead participants to more readily imagine circumstances and targets that confirm rather than disconfirm the query. This, in turn, limits researchers’ ability to understand participants’ attitudes and leads to, widely observed, lack of association between answers in trust and trustworthiness surveys and behaviour in the games in which situational context and sometimes even target is, at least partially, known.

The claim, therefore, that attitudinal measures of trust are unhelpful when it comes to providing real insight into people’s trustful and trustworthy behaviour is a significant and serious one, especially since those measures are used to gain knowledge about the state of social capital in many countries. Such information is used not only by psychologists, but also by sociologists, anthropologists, and economists. Even highly renowned sample surveys such as the General Social Survey

or the World Value Survey ask attitudinal trust questions (e.g.: “Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people?”). In the light of the described studies, in which these questions have no relation to the actual behaviour, it is important to reconsider their use or maybe change them to questions that are more specific and have proven correspondence with trustful and trustworthy actions. Even more so, since according to Yamagishi et al. (1999) “most people can be trusted” and “you can’t be too careful in dealing with people” are not on the opposite ends of a single dimension, but rather represent two separate factors.

2.2. Behavioural Measures of Trust

Another commonly used method of measuring trust and trustworthiness is through games (e.g. Trust Game, Prisoner’s Dilemma Game; Macy & Skvoretz, 1998; Yamagishi, 1999; Slemrod & Katuscak, 2002; Kiyonari et al., 2006). Nonetheless, in order to do so in line with the definition of trust accepted here, it can only be a one-shot game. When iterated, trust changes into assurance, as participants know that it is in their partner’s interest to cooperate. For instance, Axelrod (1984) indicated that even the most self-interested egoist, who typically distrusts others, would behave cooperatively in iterated prisoner’s dilemma games when faced with a partner who plays tit-for-tat.

In the current study, it was, therefore, decided to use a one-shot Trust Game. This decision was based on the broad use of this game in the measurement of trust, which would allow for comparability. Many scholars’ belief (e.g. Glaeser et al., 2000; Holm & Danielson, 2005; Kiyonari et al., 2006) that Trust Game measures trust and trustworthiness was partly supported by research showing that oxytocin (a hormone known for its role in social attachment) prevents trusting behaviour to decrease after partner’s betrayal (Baumgartner et al., 2008). This is how Camerer (2003) described why Trust Game is a good choice in trust measurement:

“Trust must be risky. Trustworthiness must also go against the Trustee’s self-interest, to test whether people are willing to sacrifice to satisfy moral obligation. Sociologists and social psychologists sometimes object that

this game “doesn’t capture all there is to trust” because the two-person one-shot game does not include the relationships, social sanctions, communication, and so many other rich features that may support or affect trust. That’s precisely the point – the game requires pure trust.” (Camerer, 2003, p. 85)

2.3. Indirect Method of Trust Measurement – Social Desirability

While currently used trust and trustworthiness measures seem ineffective in predicting trust and trustworthiness, there are several studies that suggest social desirability might work instead (e.g. DeMaio, 1984; Platow, 1994; Bekkers, 2001; Kuncel & Tellegen, 2009). Social desirability is described as a bias reflecting a tendency to enhance characteristics that are viewed as socially desirable or to minimize those that are deemed undesirable in one’s group of reference. Social desirability is believed to directly reflect societal norms and rules of conduct characteristic of a given culture and may, therefore, be related to a desire to conform or a fear of rejection. Cultural values specific for a particular society denote what is appropriate and inappropriate and set limits within which people are allowed to search for ways of behaving (DeMaio, 1984).

According to Platow (1994), social desirability may to some extent pertain to choices made in decomposed games, which are used as a measure of social value orientations. It was demonstrated that choosing a prosocial option was judged by participants as very socially desirable. Moreover, based on a large survey, Bekkers (2001) found that people who behaved in a prosocial manner were more likely to score high on a Social Desirability Scale than those who behaved in an individualistic or competitive way. More specifically, the difference was based on two subscales, one having to do with positive strategies in managing social relations and the second dealing with the tendency to abstain from profiting from others.

This led to a supposition that Social Desirability Scale may serve to identify people who are so sensitive to social rules. Such sensitivity might not only affect a tendency to respond in a socially desirable way in surveys, but also one’s actual behavior. Given that trustworthiness may be viewed as more socially desirable than trustful-

ness, current research aims to investigate the relation between Social Desirability Scale and trustworthy behaviour in the Trust Game. Since games are not always advantageous, or even possible, to be used, uncovering different measures which could serve the purpose of determining trustworthiness would be extremely beneficial.

3. Trust and Trustworthiness Across Nations

The second aim of the current research is the replication of the results obtained by various scholars (e.g. Fershtman & Gneezy, 2001; Foddy, Platow, & Yamagishi, 2009; Gleaser et al., 2000) that group membership, be it nationality, ethnicity, race, or even different university, is a powerful predictor of the degrees of trust, namely, that trust between partners decreases when they differ in group membership. Gleaser et al. (2000), for example, found that in Trust Game people returned much less when dealing with partners from another country, however the amount sent was not significantly different. In the current study, the in-group/out-group affiliation is manipulated by providing information about Trust Game partner’s nationality. Nationalities were chosen on the basis of studies by Kofta & Sedek (2005), who identified three nationalities that were most often associated with out-group prejudice in Poland. These are, in order from the strongest to the weakest, Jews, Germans, and Russians. Members of these nationalities are viewed unfavourably and with considerable suspicion, with much of negative feeling being due to conspiracy beliefs that are especially strong when it comes to Jews. Hence, in the current study, the experimental manipulation involved information about nationality of the other person in the Trust Game (Polish, Jewish, German). In the control condition, no information whatsoever about the partner was given.

Based on the previous results, it was expected that Trust, Caution, and Honesty Scales would fail to predict the Amount Sent and Returned in the Trust Game. However, higher scores on the questions pertaining to Past Trusting Behaviour and Past Trustworthy Behaviour would predict the Amount Sent (AS) and Amount Returned (AR) in the Trust Game, respectively. Additionally, it was suspected that

subjects scoring higher on the Social Desirability Scale would return more (AR) in the Trust Game. It was also expected that trust and trustworthiness would be weaker and, thus, the amounts sent and received would be lower for Trust Game partners of different nationalities than for those of Polish nationality.

4. Method

4.1. Participants

Data was collected from 156 students (94 males; 62 females) of several universities in Warsaw. The age of the participants ranged from 19 to 43 ($M = 22.36$, $SD = 2.87$).

4.2. Materials

Trust Questionnaire. For the purpose of this study, three scales from the Trust Questionnaire (Yamagishi & Yamagishi, 1994) were translated from English to Polish. The scales include the General Trust Scale (six items, $\alpha = 0.76$), the Caution Scale (six items, $\alpha = 0.62$), and the Honesty Scale (4 items, $\alpha = 0.57$). Responses to the items were to be indicated on the four-point Likert Scale (Strongly Disagree, Rather Disagree, Rather Agree, Strongly Agree).

Previous behaviour. Two scales were used to measure previous trusting and trustworthy behaviour. The purpose of using these scales was mostly exploratory, as no such coherent and reliable scales exist. Past Trusting Behaviour Scale consisted of 3 questions ($\alpha = 0.52$). Two of them (“How many times last year did you lend someone money” and “How many times last year did you lend someone your possessions”) were taken from the questionnaire used by Glaeser et al. (2000). In order to check whether the scale can achieve better discriminative power, one more question was added (“How many times last year did you buy anything on an internet site, e.g. Allegro or E-bay”). Past Trustworthy Behaviour Scale consisted of 5 questions ($\alpha = 0.59$; “How many times last year did you break a promise given to someone,” “How many times last year did you fail to meet the deadline of an assigned work,” “How many times last year did you reveal an entrusted secret?,” “How many times last year did you gossip about someone you know,” “How many times last year did you

forget to give back the money or possession you borrowed?”). Answers to both scales were to be indicated on a 7-point scale (not once, 1–2 times, 3–4 times, 5–6 times, 7–8 times, 9–10 times, more than 10 times).

Social desirability scale. The 29-item scale developed by Jolanta Wilczyńska and Radosław Drwal (1995) was used. Instead of the original true-false response key, a four-point Likert scale, identical to the one for the trust questionnaire, was used. The reliability coefficient was not, however, appreciably affected by this modification and equalled $\alpha = 0.75$.

Trust game. The classic version of the Trust Game was used (Berg et al., 1995). The amount sent by the allocator was viewed as a measure of trust, while the amount returned by the recipient was considered a measure of trustworthiness. In the first stage of the game, when all subjects were allocators, after being provided with an instruction for the game, the following questions were asked: G1: “What is the amount you give to your partner,” G2: “What is the amount you think you would be given by your partner, if he was in your place.” In the second stage of the game, when all subjects were recipients, after being provided with instructions and informed about the amount given to them by their partners (50 PLN in all cases), the following questions were asked: G3: “What is the amount that you return to your partner (after it was tripled)” and G4: “What is the amount you think would be returned by your partner, if he was in your place (after it was tripled)”.

4.3. Procedure

The participants were recruited to the study in university corridors. After signing an informed consent form, each participant was asked to provide basic demographic data (age, gender, faculty at the university). First, all participants were given the same two-part questionnaire. The first part consisted of Social Desirability, Trustworthiness, Caution, and General Trust Scales, while the second part was formed out of Past Trusting Behaviour and Past Trustworthy Behaviour Scales.

Upon completion of the questionnaire, participants were introduced to the second part of the study, the Trust Game. They were informed that the game would be played with another student who had

already provided his/her answers in the first stage of this study. Subsequently, subjects were randomly assigned to one of the four conditions and provided with information about their Trust Game partner. The information differed in terms of the nationality and the name of the partner: 1) Anna Kamińska/Andrzej Kamiński, Poland, 2) Gisela Lechner/Clemens Lechner, Germany 3) Sara Begin/Shlomo Begin, Israel 4) control group with no information about the partner whatsoever. The gender of the other person was always the same as the gender of the participant. Additionally, the information about the other person (except for the control condition) included their age (always 22 years old), sex (the same as the subject's), and the occupation (always a 4th year psychology student; in the case of partners whose nationalities were other than Polish, it was indicated that it was an Erasmus student).

The game consisted of two parts. In the first part, all participants were allocators, while in the second, all were recipients. Before each part, participants were provided with an instruction on the game. In the role of an allocator, participants were asked to fill in the blank space with the appropriate amount given to the other person (any sum out of 100 PLN) and place that information on an envelope. In the role of the recipient, participants were given an envelope with the allocation of the second person (always 50 PLN) and were asked to write how much they return to that person (any sum out of 150 PLN). The amount sent by the allocator is viewed as a measure of *trust*, while the amount returned by the recipient is considered a measure of *trustworthiness*. At the end, subjects were thanked for their participation, briefed, and given an opportunity to provide their e-mail addresses in order to be informed about the results of the study.

5. Results

Trust Questionnaire and Trust Game. In order to analyse the relationship between behaviour in the Trust Game and answers on the trust questionnaire, two multiple regression analyses were performed, separately for the Amount Sent and the Amount Returned as dependent variables, with results on General Trust, Cautiousness, and Honesty scales included as predic-

tors. A higher score on the scales indicated, correspondingly, higher general trust, cautiousness, and honesty. Both regression models proved to be non-significant, $F(3,152) = 1.473, p = .224$ for the Amount Sent and $F(3,152) = 2.051, p = .109$ for the Amount Returned. None of the included variables predicted the Amount Sent. For the Amount Returned, only the cautiousness score reached the conventional level of significance ($\beta = -.19, p = .038$).

Past Behaviour Scales. First, the correlations between the Past Behaviour Scales and other scales were computed. Higher results on the Past Behaviour Scales indicate lower declared trustworthiness and higher trusting. Past Trustworthy Behaviour Scale was significantly correlated with Social Desirability Scale ($r = -0.54, p < 0.001$) and Honesty Scale ($r = -0.26, p < 0.01$). Past Trusting Behaviour Scale, however, was not significantly correlated with any of the other scales.

To check whether the Past Behaviour Scales predicted behaviour in the Trust Game, two regression analyses were performed, separately for the Amount Sent and the Amount Returned as the dependent variables. Past trusting and past trustworthy behaviour scores were included as predictors. Neither of these models were significant, $F(2,147) = .482, p = .619$ for the Amount Sent and $F(2,147) = 1.353, p = .262$ for the Amount Returned.

However, when the relationship between individual past behaviour questions and the behaviour in the Trust Game was analysed, the question: „How many times last year did you gossip about someone you know?” proved to be significantly correlated with trustworthy behaviour in the game (G3; $r = -.23, p = .005$), as well as beliefs of other person's trustworthiness in the game (G4; $r = -.21, p = .009$). In short, people who admitted having gossiped more, returned less money to their partners and predicted their partners would also return less to them.

Social Desirability Scale and Trust Game. Further, the relationship between answers to Social Desirability Scale and behaviour in the Trust Game was investigated. For this scale, higher scores indicated higher Social Desirability. Results of the Social Desirability Scale significantly predicted the Amount Returned ($\beta = .23$), $F(1,154) = 8.275, p = 0.005$, as tested with the regression

analysis. The scale was not, however, related to the expected trustworthiness of the other person (G4), $F(1,154) = 2.432$, $p = 0.121$. Answers to Social Desirability Scale were also not related to the Amount Sent, $F(1,154) = .434$, $p = 0.511$.

Bekker (2001) reported that items which were the most accurate in predicting prosociality in decomposed games were the ones referring to positive strategies in management of social relations and to abstaining from taking advantage of others. In order to see if the same was true for trustworthiness, all items were correlated with the Amount Returned. The results confirmed the expectation based on Bekker's findings – only those items which addressed social relations proved to be related to the Amount Returned in the Trust Game. The significant effects are detailed in Table 1.

Table 1. Social Desirability items significantly correlated with Amount Returned

Question	<i>r</i>	<i>p</i>
People who ask me for favour sometimes annoy me	-.17	.040
I never deliberately said something to hurt another person	.21	.009
I am always nice, even towards people who are unkind	.16	.050
I do not always tell the truth	-.20	.015
There were instances when I cheated someone	-.21	.009
When I receive a letter, I always respond immediately	.20	.014

Gender Differences. In order to investigate the gender differences in trusting and trustworthy behaviour, two Univariate ANOVAs were conducted with gender as an independent variable and Amount Allocated and Amount Returned as dependent variables. Females ($M = 83.95$; $SD = 34.93$) tended to return significantly more than males ($M = 71.93$; $SD = 33.14$), $F(1, 154) = 4.711$, $p = .031$, $\eta^2 = .030$. There were no gender differences in terms of the Amount Sent (Females: $M = 57.50$; $SD = 28.28$; Males: $M = 54.09$; $SD = 30.47$), $F(1, 154) = .495$, $p = .483$.

Group Membership. The effects of the group membership were tested using the

Univariate ANOVAs with other person's Nationality as an independent variable and the game measures as dependent variables. There was a significant difference in the Amount Sent, $F(3,152) = 4.604$, $p < 0.01$, $\eta^2 = .083$. Post-hoc analysis (Bonferroni) revealed significant differences between the Control and the Polish partner groups (95% CI for difference [-35.55, -1.12]), as well as between the Control and the German partner groups (95% CI for difference [-39.28, -4.86]). The difference between the Control and the Israeli partner groups was not significant. Means for this analysis are in Table 2.

Table 2. Mean Amount Sent according to Partner's Nationality

Partner's Nationality	M	SD
Polish	60.08	25.83
German	63.82	27.17
Israeli	57.18	31.01
Control group	41.74	29.89

There also was a significant effect for the Amount Returned, $F(3, 152) = 3.991$, $p = 0.009$, $h^2 = .073$. Post-hoc analysis (Bonferroni) showed significant differences between the Control and the Polish partner groups (95% CI for difference [-43.17, -3.06]), as well as the Control and the Israeli partner groups (95% CI for difference [-40.69, -8.56]). The difference between the Control and the German partner groups was not significant. Means for this analysis are detailed in Table 3.

Table 3. Mean Amount Returned by Partner's Nationality

Partner's Nationality	M	SD
Polish	86.32	35.14
German	74.21	31.20
Israeli	83.97	39.46
Control group	63.20	26.35

Summing up, the study did not replicate the findings that the information about partner's out-group membership results in lower trusting and trustworthy behaviour in the Trust Game. Rather, any informa-

tion about the partner resulted in higher allocations and returns, in comparison to the control condition.

6. Discussion

The results of the current study confirm the result obtained by other researchers (Glaeser et al., 2000; Holm & Danielson, 2005) that commonly used attitudinal measures of trust and trustworthiness do not predict the actual behaviour in games. Contrary to the results obtained by Glaeser et al. (2000), measures of past behaviour were also unrelated to the actions in the game. Nevertheless, the fact that one question (“How many times last year did you gossip about someone you know”) predicts trustworthiness quite well is encouraging. It might mean that past behaviour scale, if constructed out of such highly prognostic items may prove to be valid. It would be, therefore, advantageous, to undertake an effort to formulate and analyse a large number of such past behaviour items, some of which might serve well in predicting trustful and trustworthy behaviour.

Unfortunately, no measure correlated with trustful behaviour in the current study. The fact that trustworthiness was related to at least some variables (gender, social desirability, gossiping), while trustfulness was completely independent, might mean that trustworthiness is a more stable personal characteristic, whereas trust depends more on situational cues. It is however possible that, since trust entails relinquishing some degree of control in favour of another person, a measure composed of questions relating to people’s willingness to do just that could be devised. The trust surveys used so far ask very general questions (e.g. “Most people can be trusted,” “Most people are basically good and kind,” etc.) which create problems in terms of analysis of participants’ responses, since it is questionable what was on their mind. When answering questions about “most people,” did they think about an average of population? People they know? Or maybe the last time they were cheated by a salesclerk? Such uncertainty might lead to variability in responses that cannot be controlled.

Consequently, it could be better to devise a survey consisting of questions that refer to possibility of trusting others in specific tasks and/or situations. Quite impor-

tantly, these questions would also need to indicate, at least in a general sense, the object of trust, e.g. member of a family, friend, work colleague, stranger, etc. A survey so created would fulfill the conditions of a good attitudinal measure proposed by Ajzen and Fishbein (1977), namely, the need for a defined target, action, and context, or at least two of these. They suggested that instruments having high correspondence between predictor and criterion produce strong relationships between attitudes they measure and actual behaviours.

An additional problem with measures of trust and trustworthiness employed in the current study refers to the fact that they were a directly translated form of Yamagishi’s (1994) questionnaire used in the United States. Although reliability coefficients were comparable with the original version, it might be the case that cultural differences make this questionnaire not valid in Polish circumstances. Toshio and Midori Yamagishi chose the best items out of a large amount of them, while in the current research only those from the final questionnaire were taken into consideration. This might have led to a decreased validity of the measure, especially since some researchers (e.g. Holms & Danielson, 2005; Henrich et al., 2005) argue that trust may have different facets across cultures.

6.1. Social Desirability

It is encouraging, however, that the present research confirmed the hypothesis that the Social Desirability Scale is accurate in predicting trustworthy behaviour. It might be that the people who express high need for approval have a tendency not only to respond in a socially desirable manner in questionnaires but also to behave in such a way. Although the Trust Game in the current research is still a highly constrained, experimental situation with limited external validity, it may be argued that, since being trustworthy is not only highly socially desirable but also refers almost exclusively to actions in social settings, people with heightened need for social approval could have a tendency to act in a similar way also in their daily lives. It is suggested that social desirability might refer to people’s inclination to look good not only in the eyes of an experimenter but also in the eyes of other members of society. Moreover, trustworthiness can be thought of as a highly socially

desirable trait and a key element of social norms of fairness (cf. Platow, 2001), hence providing strong inducement for those for whom self-presentation is crucial to act in such a way.

This supposition is additionally confirmed by the results obtained by Bekkers (2001), who found that people who acted more cooperatively or altruistically in decomposed games scored higher on Social Desirability Scale, but only on items referring to interpersonal conflict or concern. Similar results were obtained in the current study. Analyses revealed that the particular Social Desirability Scale questions which distinguished trustworthy from untrustworthy subjects were the ones referring to interpersonal relations. Items relating to personal achievement had no predictive power for trustworthiness whatsoever.

6.2. Gender Differences

In the current study, females behaved in a more trustworthy manner, indicated by greater amounts of money returned to their partners in the Trust Game. This effect was not surprising and was already documented in the previous studies (Croson & Buchan, 1999; Dollar, Fisman, & Gatti, 2001). There were, however, no gender differences in trust indicated by behaviour in the game, which is in line with the results of Croson & Buchan (1999), but not those of Gleaser et al. (2000). The previously reported gender differences in trust could have been related to greater risk aversion among women (Arano, Parker, & Rory, 2010; Borghans et al., 2009). If this was the case, the game situation in the current study was likely not experienced as risky, resulting in no differences between males and females.

6.3. Trust and Trustworthiness Across Nations

The results obtained in the current study failed to replicate the previous findings that trust and trustworthiness are negatively affected by information about partner's out-group membership (Fershtman & Gneezy, 2001; Foddy, Platow, & Yamagishi, 2009; Gleaser et al., 2000). This lack of replication might have been caused by extensiveness of the information provided about the partners. It consisted of not only information on their nationality, but also their name, age, gender, and major – some

of which provide individuating information about the other person. According to the Social Identity model of Deindividuation Effects (SIDA, Reicher, Spears, & Postmes, 1995), expectations of others' trustworthiness may be based on either *cues to personal identity* or *shared group membership*. When cues to personal identity are provided, e.g. name or picture, detrimental effect of distinct group membership is diminished and perceived trustworthiness of a partner increases (Lee, 2008; Tanis & Postmes, 2005). This effect, however, works only for the out-group members and can even have a disadvantageous outcome for the in-group members, because it underlines partner's uniqueness and draws attention away from that person's in-group membership. In line with this reasoning, research by Ritov & Kogut (2011) showed that generosity towards the members of the out-group increased if personalizing information about that person was provided (e.g. through name or picture).

The current study does not answer the question on the extent to which the two types of information – cues to social or personal identity – separately affected the obtained results. In order to gather such information, an experiment with two independent variables (group membership x cues to personal identity) would need to be conducted.

6.4. Limitations of the Study

It is also important to mention other two possible limitations of the present research. First of all, some researchers believe that the use of hypothetical money decreases variability of the results (e.g. Shaw, 1976). Although studies exist (e.g. Gillis & Hettler, 2007) that suggest no significant differences between the use of imaginary and real money and despite the fact that, in the current study, participants were explicitly asked to visualize their behaviour is the real situation, it still might have made them more carefree in their allocations and returns.

The second limitation is related to the artificiality of the manipulation of partner's nationality. Because of the emphasis put nowadays in Poland on the importance of good cross-cultural relations and the evil of anti-Semitism, participants might have guessed the purpose of the study and behaved differently than they would in real-life settings. It was also suggested

by Dovidio et al. (2002) that when a society moves away from an open expression of prejudice, it might become more covert and even unconscious. Our manipulation, in which nationality was explicitly stated, might have sensitized participants to their prejudices and made them more careful not to express them. Hence, future studies should consider employing a less artificial way of manipulating partner's group membership.

7. Conclusions

Concluding, the present study shed a new light on two areas related to trust. First of all, it proved inadequacy of standard attitudinal measures of trust and trustworthiness. It was found that Social Desirability Scale might serve better in the assessment of trustworthiness than either traditionally used questionnaires or measurements of past behaviour. Unfortunately, no measures were uncovered which could serve as the predictors of trusting behaviour. Possibilities of constructing a valid trust questionnaire are, however, suggested in the discussion. Secondly, the study failed to replicate the results that information about partner's out-group membership decreases trust and trustworthiness in the Trust Game. The reason for such results could stem from providing participants with information about partner's personal identity, which might have annulled the detrimental effects of the out-group membership.

The aim of a further study would be to enhance the external validity of the results by providing more real-life experimental settings, especially when it comes to the nationality manipulation, as well as to reconsider the use of hypothetical money. Moreover, a study could be performed in which presence of cues to both social and personal identity was manipulated. Such a manipulation would make the effects observed in the current study clearer.

Footnote

1 Acknowledgements: I thank Professor Eugene Burnstein for his insightful suggestions during the development of this study, as well as his invaluable comments on the text.

Funding: This research was supported by National Science Centre grant (DEC-2012/07/N/HS6/02252)

References

- Ajzen, I. and Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84, 888–918, <http://dx.doi.org/10.1037/0033-2909.84.5.888>.
- Arano, K., Parker, C. and Rory, T. (2010). Gender-based Risk Aversion and Retirement Asset Allocation. *Economic Inquiry*, 48(1), 147–155, <http://dx.doi.org/10.1111/j.1465-7295.2008.00201.x>.
- Axelrod, R. (1984). *The Evolution of Cooperation*, New York: Basic Books, [http://dx.doi.org/10.1016/0140-1750\(86\)90032-1](http://dx.doi.org/10.1016/0140-1750(86)90032-1).
- Baier, A.C. (1986). *Trust and Antitrust*, Ethics, 96, <http://dx.doi.org/10.1086/292745>.
- Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., Fehr, E. (2008). Oxytocin Shapes the Neural Circuitry of Trust and Trust Adaptation in Humans. *Neuron*, 58(4), 639–650, <http://dx.doi.org/10.1016/j.neuron.2008.04.009>.
- Bekkers, R. (2001). The Social Desirability of Social Value Orientations. *Report written during a research visit to the Department of Social Psychology, Free University, Amsterdam*. Retrieved July 13, 2010, from www.fss.uu.nl/soc/homes/bekkers/sdresnot.pdf.
- Berg, J., Dickhaut, J. and McCabe, K.A. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behaviour*, 10(1), 122–142, <http://dx.doi.org/10.1006/game.1995.1027>.
- Borghans, L., Heckman, J.J., Golsteyn, B.H.H. and Meijers, H. (2009). Gender Differences in Risk Aversion and Ambiguity Aversion. *Journal of European Economic Association*, 7(2/3), 649–658, <http://dx.doi.org/10.3386/w14713>.
- Caldwell, C. and Clapham, S.E. (2003). Organizational trustworthiness: An international perspective. *Journal of Business Ethics*, 47, 349–358.
- Camerer, C. (2003). *Behavioural Game Theory. Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press, <http://dx.doi.org/10.1007/s11615-004-0067-y>.
- Croson, R. and Buchan, N. (1999). Gender and Culture: International Experimental Evidence from Trust Games. *American Economic Review*, 89(2), 251–279, <http://dx.doi.org/10.1257/aer.89.2.386>.
- DeMaio, T.J. (1984). Social Desirability and Survey Measurement: A Review. In: C.F. Turner and E. Martin (eds.), *Surveying Subjective Phenomena* (Vol. 2, pp. 257–281). New York: Russell Sage.
- Dollar, D., Fisman, R. and Gatti, R. (2001). Are Women Really the 'Fairer' Sex? Corruption and Women in Government. *Journal of Economic Behaviour and Organization*, 46(4), 423–429, [http://dx.doi.org/10.1016/S0167-2681\(01\)00169-x](http://dx.doi.org/10.1016/S0167-2681(01)00169-x).

- Dovidio, J.F., Gaertner, S.L., Kawakami, K. and Hodson, G. (2002). Why Can't We Just Get Along? Interpersonal Biases and Interracial Distrust. *Cultural Diversity and Ethnic Minority Psychology*, 8(2), 88–102, <http://dx.doi.org/10.1037/1099-9809.8.2.88>.
- Wilczyńska, J. and Drwal, R.L. (1995). Opracowanie Kwestionariusza Aprobata Społecznej (KAS) J. Wilczyńskiej i R.L. Drwala. In: R.L. Drwal (ed.) *Adaptacja kwestionariuszy osobowości. Wybrane zagadnienia i techniki* (s. 57–66). Warszawa: Wydawnictwo Naukowe PWN.
- Fershtman, C. and Gneezy, U. (2001). Discrimination in a Segmented Society: An Experimental Approach. *Quarterly Journal of Economics*, 116(1), 351–377, <http://dx.doi.org/10.1162/003355301556338>.
- Foddy, M., Platow, M.J. and Yamagishi, T. (2009). "Group-based trust in strangers: The role of stereotypes and expectations". *Psychological Science* 20, 419–422, <http://dx.doi.org/10.1111/j.1467-9280.2009.02312.x>.
- Gillis, M.T. and Hettler, P.L. (2007). Hypothetical and Real Incentives in the Ultimatum Game and Andreoni's Public Goods Game: an Experimental Study. *Eastern Economic Journal*, 33(4), 491–510, <http://dx.doi.org/10.1057/ej.2007.37>.
- Glaeser, E.L., Laibson, D.I., Scheinkman, J.A. and Soutter, C.L. (2000). Measuring Trust. *The Quarterly Journal of Economics*, 115(3), 811–846, <http://dx.doi.org/10.1162/003355300554926>.
- Hardin, R. (2006). *Trust*. Cambridge: Polity Press.
- Hayashi, N. and Yamagishi, T. (1998). Selective Play: Choosing Partners in an Uncertain World. *Personality and Social Psychology Review*, 2(4), 276–289, http://dx.doi.org/10.1207/s15327957pspr0204_4.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Hill, K., Gil-White, F., Gurven, M., Marlowe, F., Patton, J. Q., Smith, N., & Tracer, D. (2005). Economic man in cross-cultural perspective: behavioural experiments in 15 small-scale societies. *Behavioural and Brain Sciences*, 28(6), 795–815, <http://dx.doi.org/10.1017/s0140525x05000142>.
- Holm, H.J. and Danielson, A. (2005). Tropic Trust versus Nordic Trust: Experimental Evidence from Tanzania and Sweden. *The Economic Journal*, 115(4), 505–532, <http://dx.doi.org/10.1111/j.1468-0297.2005.00998.x>.
- Kiyonari, T., Yamagishi, T., Cook, K. S. and Cheshire, C. (2006). Does Trust Beget Trustworthiness? Trust and Trustworthiness in Two Games and Two Cultures: A Research Note. *Social Psychology Quarterly*, 69(3), 270–283, <http://dx.doi.org/10.1177/019027250606900304>.
- Knack, S. and Keefer, P. (1997). Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *Quarterly Journal of Economics*, 112(4), 1251–1288, <http://dx.doi.org/10.1162/003355300555475>.
- Kofta, M. and Sedek, G. (2005). Conspiracy Stereotypes of Jews During Systemic Transformation in Poland. *International Journal of Sociology*, 35(1), 40–64.
- Kohn, M. (2008). *Trust. Self-Interest and the Common Good*. Oxford: University Press.
- Kuncel, N.R. and Tellegen, A. (2009). A Conceptual and Empirical Re-examination of the Measurement of the Social Desirability of Items: Implications for Detecting Desirable Response Style and Scale Development. *Personnel Psychology*, 62, 201–228, <http://dx.doi.org/10.1111/j.1744-6570.2009.01136.x>.
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. W. (1997). Trust in Large Organizations. *American Economic Review Papers and Proceedings*, 87(5), 333–338, <http://dx.doi.org/10.3386/w5864>.
- Lee, E. (2008). When Are Strong Arguments Stronger Than Weak Arguments? Deindividuation Effects on Message Elaboration in CMC. *Paper presented at the annual meeting of the International Communication Association, TBA, Montreal, Quebec, Canada*. Retrieved July 12, 2010, from http://www.allacademic.com/meta/p229883_index.html.
- Leslie, C.R. (2004). *Trust, Distrust, and Antitrust*, Texas Law Review, 3(82).
- Macy, M.W. and Skvoretz, J. (1998). The Evolution of Trust and Cooperation Between Strangers: A Computational Model. *American Sociological Review*, 63(5), 638–660, <http://dx.doi.org/10.2307/2657332>.
- Platow, M.J. (1994). An Evaluation of the Social Desirability of Prosocial Self-Other Allocation Choices. *The Journal of Social Psychology*, 134(1), 61–68, <http://dx.doi.org/10.1080/00224545.1994.9710884>.
- Reicher, S.D., Spears, R. and Postmes, T. (1995). A social identity model of deindividuation phenomena. In W. Strobe and M. Hewstone (eds.), *European review of social psychology* (pp. 161–198). Chichester, UK: Wiley, <http://dx.doi.org/10.1080/14792779443000049>.
- Ritov, I., and Kogut, T. (2011). Ally or adversary: the effect of identifiability in inter-group conflict situations, *Organizational Behavior and Human Decision Processes*, 116(1), 96–103, <http://dx.doi.org/10.1016/j.obhdp.2011.05.005>.
- Rousseau, D.M., Sitkin, S.B., Burt, R.S. and Camerer, C. (1998). Not So Different After All: A Cross-discipline View of Trust. *Academy of Management Review*, 23(3), 393–405, <http://dx.doi.org/10.5465/amr.1998.926617>.
- Shaw, J.I. (1976). Response-Contingent Payoffs and Cooperative Behaviour in the Prisoner's Dilemma

- Game. *Journal of Personality and Social Psychology*, 34(5), 1024–1033, <http://dx.doi.org/10.1037/e668292012-312>.
- Slemrod, J.B. and Katuscak, P. (2005). Do Trust and Trustworthiness Pay Off? *Journal of Human Resources*, 40(3), 621–646, <http://dx.doi.org/10.3368/jhr.xl.3.621>.
- Sztompka, P. (2007). *Zaufanie. Fundament społeczeństwa*. Kraków: Znak.
- Tanis, M. and Postmes, T. (2005). A Social Identity Approach to Trust: Interpersonal Perception, Group Membership and Trusting Behaviour. *European Journal of Social Psychology*, 35(3), 413–424, <http://dx.doi.org/10.1002/ejsp.256>.
- Yamagishi, T. and Yamagishi, M. (1994). Trust and Commitment in the United States and Japan. *Motivation and Emotion*, 18(2), 129–166, <http://dx.doi.org/10.1007/bf02249397>.
- Yamagishi, T., Kikuchi, M. and Kosugi, M. (1999). Trust, Gullibility, and Social Intelligence. *Asian Journal of Social Psychology*, 2(1), 145–161, <http://dx.doi.org/10.1111/1467-839x.00030>.