Data Report for 2023 SNSD Surveys in England and Italy

Ron Burt, Bocconi University and University of Chicago

This is a technical note written to provide background text for publications on data from two online surveys on social networks and social decisions (SNSD). The interviews collect social network data with a slight variation on the General Social Survey (GSS) name generator, accompanied by several name interpreters. The data are sufficient to compute network measures of size, composition, density, and access to structural holes. The interview includes four trust measures: a baseline of trust opinion as elicited by the GSS trust question, and trust behavior in three games respondents play with three other people drawn at random from a respondent pool: cooperation in a one-shot Prisoner's Dilemma Game, trust behavior in the Trust Game, and trustworthy behavior in the Trust Game. I will refer to the England and Italy surveys as the SNSD surveys. The surveys were conducted through Kantar, a market research company, and targeted national populations of college-educated adults in England (January, 2023) and Italy (April, 2023). The goal was to experiment with a rapid, low-cost alternative to familiar telephone and face-to-face surveys. There are three sections to this note: Sample, Instrument, and Interview.

SAMPLE: HETEROGENEITY, NOT REPRESENTATION

In an early project combining game experiments with population surveys, Fehr et al. (2003) offered succinct, compelling argument based on the two kinds of evidence having complementary strengths and weaknesses. Experiments provide behavioral data in a controlled setting, but often suffer from homoogeneous subject pools and self-selection biases created by who choses to be a subject. Surveys can provide representative population heterogeneity, but often limit analysis to respondent attitude as an indicator of respondent behavior. Together, surveys and experiments have the potential to provide representative sample data on behavior in a relatively controlled situation.

The SNSD surveys offer respondent heterogeneity, but they are not representative of their populations. The failing is not unique to the specific company hired to do the survey fieldwork. Broad access to the internet makes it possible for analysts to

---

Acknowledgment: I wrote this note with frequent comment from colleague Sonja Opper and occasional comment from colleague Nicolo Cavalli. The instrument in the Appendix was written by Ron Burt and Sonja Opper primarily drawing questions from instruments they had used in the US and China. The translation of English into Italian was checked by reverse translation, and readings by Nicolo Cavalli, Diego Gambetta, and Giuseppe Soda. The online questionnaire was developed by a Chicago software engineer, Joao Santos, with the support of the Booth School of Business, University of Chicago ($16,600 from Ron Burt's Chicago research funds). Fieldwork was supported by Bocconi University (7,623 euro from Nicolo Cavalli's Bocconi research funds, 9,950 euro from Ron Burt's Bocconi research funds, and 8,225 euro from a Bocconi Senior Investigator award to Sonja Opper). Total material costs were less than forty-three thousand euro.

assemble large numbers of self-selected respondents — Facebook or Twitter users who agree to respond to a question, Amazon Mechanical Turk workers who accept a task, or, in the case of the SNSD surveys, people in the Kantar respondent pool who agree to be interviewed. The non-representative point seems obvious, but non-representative sample data akin to the SNSD samples are often presented as representative in published research, so I want to be explicit about the point (briefly).

Ideally, a survey of a representative sample draws respondents with known probability from a population so that inferences about the population can be made from sample statistics. For example, the General Social Survey uses U.S. Census Bureau data to stratify the adult, noninstitutionalized population by geography, race, and income down to a neighborhood level, within which people are selected for interview by age, gender, and employment in proportion to the neighborhood's census-tract data. Any one-sentence description of the sampling process oversimplifies a complex process (cf., Smith et al., 2019: Appendix A), but the goal is that every adult, noninstitutionalized person in the population has a known probability of being interviewed. Estimates of association between variables X and Y in the sample data can be used to make inferences about the strength of association between X and Y in the population.

The national populations for the SNSD surveys are native-speaker (native English speaker for the survey in England; native Italian speaker for the survey in Italy), college-educated, adults age 30 to 65 taking the interview on a device other than a moble phone (to preserve the visibility of the network name interpreter questions). In the interest of studying work integrated with family, older residents beyond retirement are put aside and young adults 18-29 are put aside. The survey fieldwork consists of the marketing research company announcing to eligible people in their respondent pool that the interview is available (along with a note on compensation and interview duration). Interviews are completed on a first-come, first-served basis. Given a contracted total number of interviews to be provided, respondents in any specific sampling category are turned away when a sufficient number of them have completed interviews. For a 500-person sample intended to be 50% female, women would be denied entry to the interview after 250 women had completed the interview.

Consider the age of SNSD respondents. Table 1 on the next page compares the age distribution of people who completed the SNSD interview in England versus the census distribution of age in Great Britain. England is different from Great Britain, but Table 2 is sufficient for the purposes here. Seven hundred people in England completed the SNSD interview. Of those, 164 were under 40 years of age (23.4%) and 137 were over the age of 59 (19.6%). In comparison to the British population, older people are over represented in the SNSD sample, and younger people are under represented. Older people have more spare time to participate in surveys, and younger people often participate via mobile phones, which were excluded here.

If similar percentages in sample and population were a concern, then the fieldwork could have been continued to allow young people more time to enter the interview, and older people could have been denied access after their quota of 89 completions had been met (12.7% of 700 would have been 89).

Thus, proportions in a volunteer sample can be engineered to match population proportions, but it need not make the sample representative of the population. Representation is disrupted by the kinds of people who sign up to be available for interview (company population), and the kinds of people who get to the interview early so they are included in the sample (chat groups among heavy users cluster members early in noteworthy interviews).

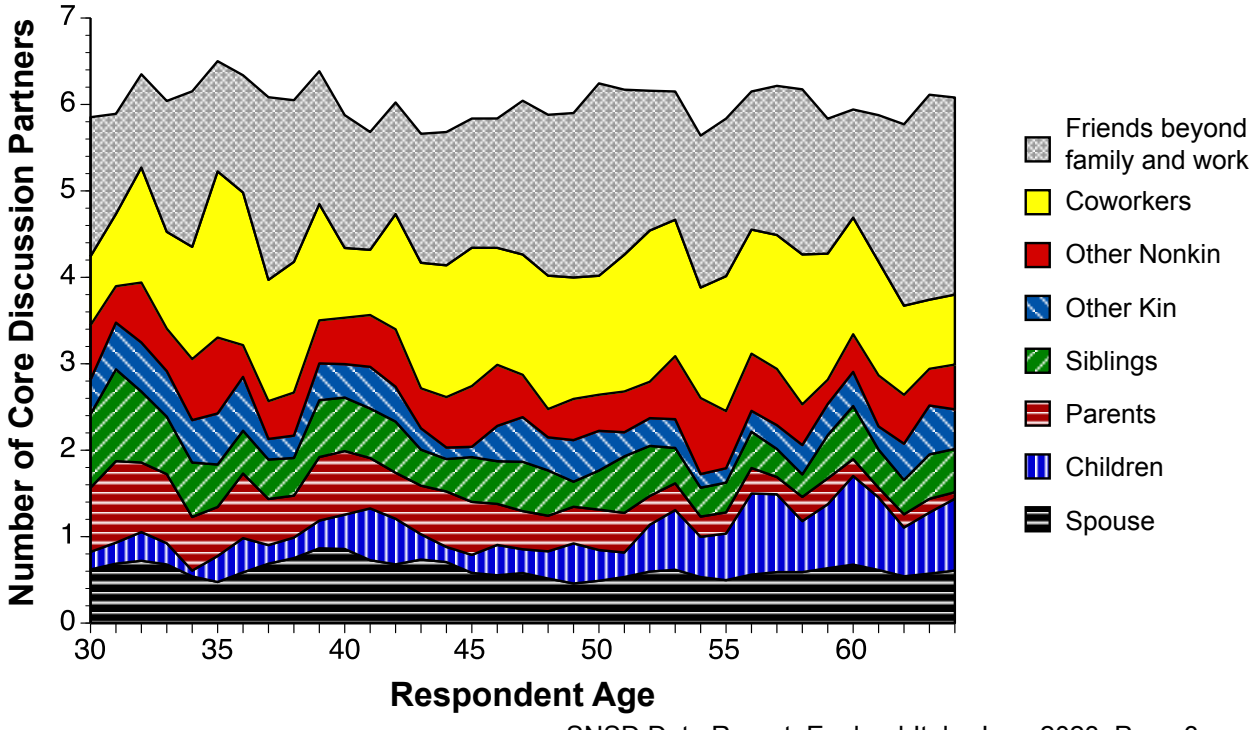
Table 1.
Age Distributions.

Age in Years	2023 SNSD	2021 Census
< 40	23.4%	29.8%
41 - 49	27.1%	27.7%
50 - 59	30.0%	29.9%
60 - 65	19.6%	12.7%
Total	100.1%	100.1%

NOTE — First column lists percent of completed interviews in each category from the SNSD survey in England. Second column is for Great Britain from the Office for National Statistics.

Consider Figure 1 below. The graphed data come from the SNSD survey in England. Respondents are ordered by age on the horizontal axis. Areas above each year show the relative number of people cited as people with whom the respondent "discusses important matters." For example, 31-year-old respondents on average cite one parent, one sibling, and 1.6 friends beyond work and family. Two of three cite their spouse or a partner with whom they are living as a spouse. I'll explain the data in a moment, but for now, notice how stable network composition is across respondent age. There is a

Figure 1.
Network Composition by Age from the SNSD Survey in England.
(Scores are moving averages across two-year windows.)



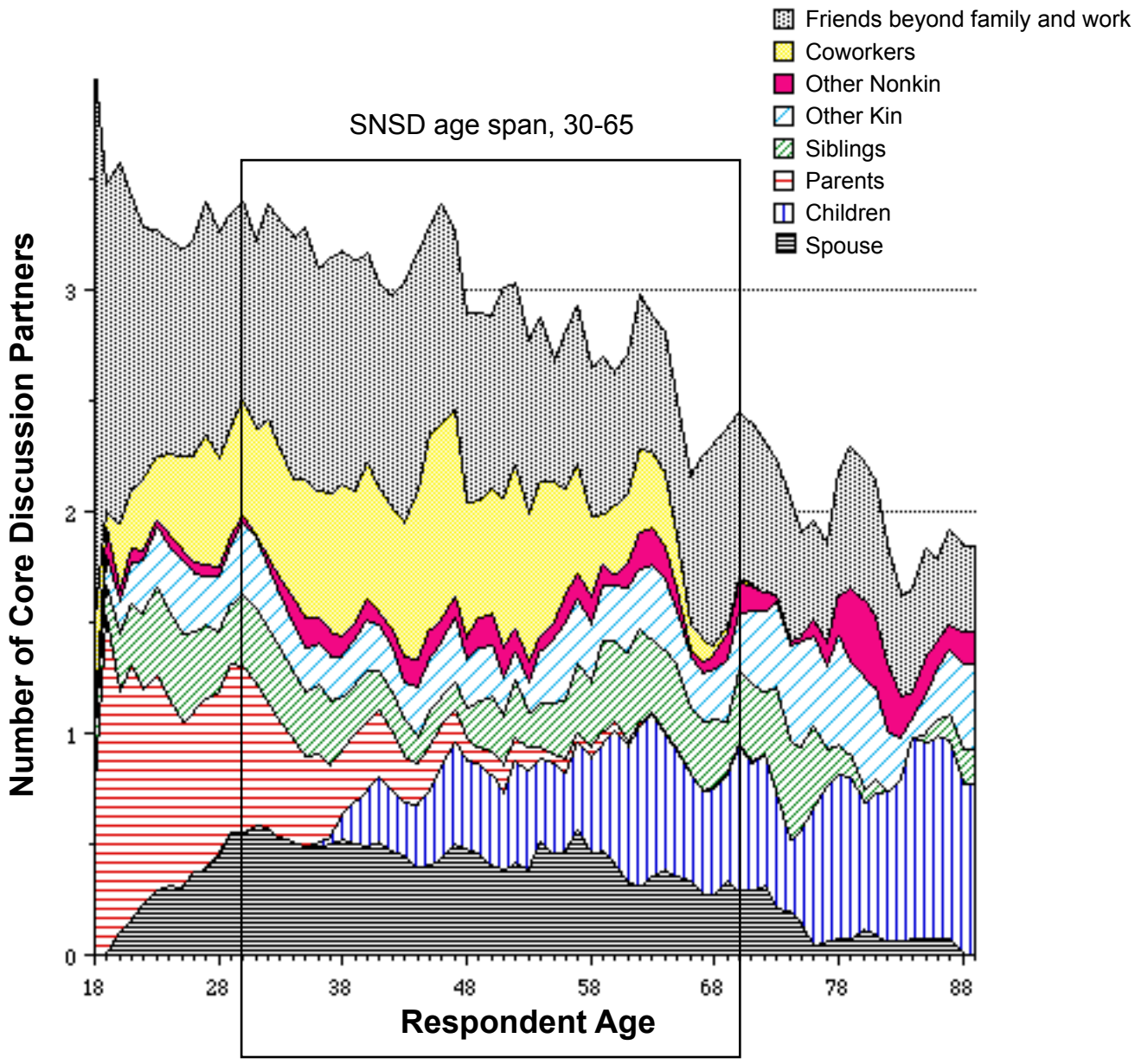


Figure 2.
Network Composition by Age in the 1985 General Social Survey (GSS).
(Scores are moving averages across two-year windows.)

transition during the 30s to the 50s in which children replace parents as discussion partners, but the top of the area graph is consistently about six discussion partners, and family members (striped areas) are consistently about three of the six.

Now compare the Figure 1 graph with the same kind of graph in Figure 2 based on General Social Survey data, data that are representative of the American population. The graph is from Burt (1991:11). The area enclosed in a box corresponds to respondent age in the SNSD surveys, 30 to 65. There are interesting life transitions apparent in the graph. For example, children replace parents as discussion partners as Americans reach their late 40s, and coworkers become much less of a presence in the networks of Americans during their 60s. But for the purposes here, the clearest pattern in Figure 2 is that older people have smaller networks. Network size is stable over age in the Figure 1 SNSD data, but the representative data in Figure 2 show

a clear decrease in size for older respondents. Data for the two figures differ by methodology and 40 years. The 1985 data for Figure 2 describe pre-internet society. It is certainly easier today for people to stay in touch with distant contacts than it was in 1985. However, having become older myself, I read Figure 2 as evidence of the withdrawal from socal activity associated with age. It takes energy to maintain a social network, and that energy is less abundant, on average, in older age.

Not for everyone. Some older people have energy to explore and maintain diverse activities. Those energetic older people are more likely to have networks larger than other people their age, and more likely to sign up with a marketing company as a volunteer to be interviewed. I assume that the older respondents in Figure 1 are disproportionately individuals in the national population who are more socially active than the average person their age in the population. The networks around those people do not shrink as is typical of people their age.

Table 1 shows there are a disproportionate number of older people in the SNSD sample, but the key point for analysis — illustrated in Figures 1 and 2 — is that the networks around those older people do not represent the networks around people their age in the national population. The SNSD older people can be studied in comparison to less socially active people, but not as representative of their age bracket in the national population. Association statistics computed from the SNSD sample data can be tested for stability and magnitude across the diversity of people interviewed, but the statistics do not indicate strength of association in the population.

INSTRUMENT

A copy of the SNSD questionnaire is attached as an appendix to this report. The copy is the English language text that was translated into Italian for the survey in Italy. The same text was used for the survey in England, except three questions were not included in the England survey (indicated below). Here is the flow of the instrument, by screen number (bottom right in each screen), then content:

1. Splash page (respondent sent here from Kantar website).
2. Informed consent.
3. Background demographics and opinion.
4. Identification as Northern or Southern Italian (only in Italy survey).
5. Four emotional resilience questions (only in Italy survey).
6. Expanded General Social Survey (GSS) network name generator.
7. Interactive editing of named discussion partners.
8. Gender name interpreter.
9. Role relation name interpreter.
10. Emotional energy name interpreter.
11. Emotional closeness name interpreter (from GSS).

12. Trust name interpreter (Response options are from GSS trust question on Screen 14, Screen 12 is only in Italy survey.).
13. GSS name interpreter eliciting relations among named discussion partners (from GSS).
14. GSS trust question.
15. Explanation of upcoming behavioral games. (“Next” button is on delay to prevent respondents moving too quickly to next page.)
16. Prisoner’s Dilemma Game (cooperate yes/no).
17. Trust Game (trustor).
18. Reflection on choice in trust game.
19. Trust Game with each cited discussion partner.
20. Trust Game (trustee).
21. Summary of game moves and respondent earnings.
22. Reflection on game behavior.
23. Sign off (respondent returned to Kantar website).

Items Other than Network and Games

The instrument begins with a splash screen and consent screen, followed by Screen 3 containing generic background questions about gender, age, marital status, country, and employment. Education is used as an indicator of socioeconomic status. Occupation involves extensive coding and reliability issues that we do not need to address in this initial work (Smith et al., 2019: Appendix F). Questions about relative success are included to provide a control for the trust-inhibiting effect of a closed network around people who feel more successful in their line of work (Burt, Opper & Holm, 2022).

Screens 4 and 5 are only in the Italy survey. Origin questions on Screen 4 are to link the SNSD survey data with research on open, cosmopolitan networks in the north versus networks closed around family in the south. The sum of the four items on Screen 5 is an index of resilience to emotional stress. Two items — “able to adapt to change” and “tend to bounce back after illness or hardship” — are recommended by Vaishnavi, Connor, and Davidson (2007) as a two-item version of the 25-item Connor-Davidson (2003:78) resilience scale. The other two items on Screen 5 are adapted from the item reported by Connor and Davidson (2003:80) to have the highest item-total correlation (.70 correlation, “think of self as strong person”), and an item with high loading on the first factor of the 25 resilience items (also page 80, .64 factor loading, “when things look hopeless, I don’t give up”).

Network Name Generator

The network items are context sensitive (see Fischer, 2009; Paik & Sanhagrin, 2013, regarding the 2004 GSS network data) and less exciting than the behavioral games, so the network items appear in the interview before the games to avoid confounding the network data with the more exciting behavioral games.

The core discussion network around each respondent is elicited using the name generator and interpreter method adopted in the 1985 General Social Survey, GSS (Burt, 1984; Marsden, 1987, 2011; Perry et al., 2018; Smith et al., 2019: 572 ff.; Marsden, Fekete & Baum, 2021). The respondent is asked to name individuals with whom he or she has a kind of relationship (name generator question), then the respondent is asked questions about each individual named (name interpreter questions). The name generator on Screen 6 is the GSS ”discuss important matters” generator: “From time to time, most people discuss important matters with other people. … Please enter to the right nicknames for the 6 or 7 people with whom you recall discussing matters important to you during the last six months.” Consistent with GSS protocol, Screen 7 serves to remove duplicate names (if there are any), and encourage the respondent to name additional contacts. .

(1) **Confidentiality**. The SNSD name generator is familiar from the GSS, but three variations warrant note. The first is confidentiality. Following the GSS, the name generator explicitly says not to use full names. To preserve respondent privacy, the GSS generator instructs respondents: “Just tell me their first names or initials.” The SNSD instrument goes a step further and ask just for nicknames and initials. Contact names are not used in the data analysis, but they are essential in the interview because the respondent needs to be able to distinguish individual contacts in the next few questions. Taking one step further, no names are recorded in the SNSD surveys. The software only records names by their citation order (first person named, second person named, etc.). No actual names are recorded. An element of trust is involved here. The respondent is told on Screen 6 that no names will be recorded, but that statement can be expected to not be as convincing as it is in a face-to-face interview in which the respondent does not give names to the interviewer (e.g., Burt, et al., 2022),

(2) **Excessive Consent**. The double consent warrants note. Staff in the market research company hired to recruit respondents were familiar with brief survey opinion items, but not survey network questions. Despite a Bocconi-approved EU consent form (Screen 2), and despite network confidentiality assurances (no cited names recorded, and no knowledge of respondents other than the demographics on Screen 3), Kantar staff insisted on a second consent on the name generator page (see the middle of Screen 6).

The second consent seems to have made respondents nervous. As noted below, the second consent was the primary place in the interview where respondents quit the interview. Few quit after the second consent (35 of 1174 English entrants to the interview, 29 of 925 Italian entrants, see Tables 5 and 6 in next section). The second consent adds addition sampline bias to the bias discussed in the previous section, in particular obscuring single, unemployed respondents (see box on the next page).

(3) **Network Size**. A third variation from the GSS is to ask for more names. The SNSD name generator asks for 6 or 7 names, up to 8. The GSS name generator stops at 5 names. Five is a sufficient number of names if many respondents provide them (Merluzzi and Burt, 2013), but three is the average number of people cited in the GSS network data. Analysis of the GSS network data (Burt, 1986) show that

family members are likely to be the first and second people cited. Contacts beyond the family are likely as persons named later.

Without losing data on key family contacts, a goal here was to secure data on close contacts beyond the family in order to resolve debate over trust in strangers lowered by strong family ties versus closed networks more generally (see Burt, 2021, for review, with key empirical results in Ermisch & Gambetta, 2010; Aassve, Conzo & Mattioli, 2021; Burt et al., 2022).

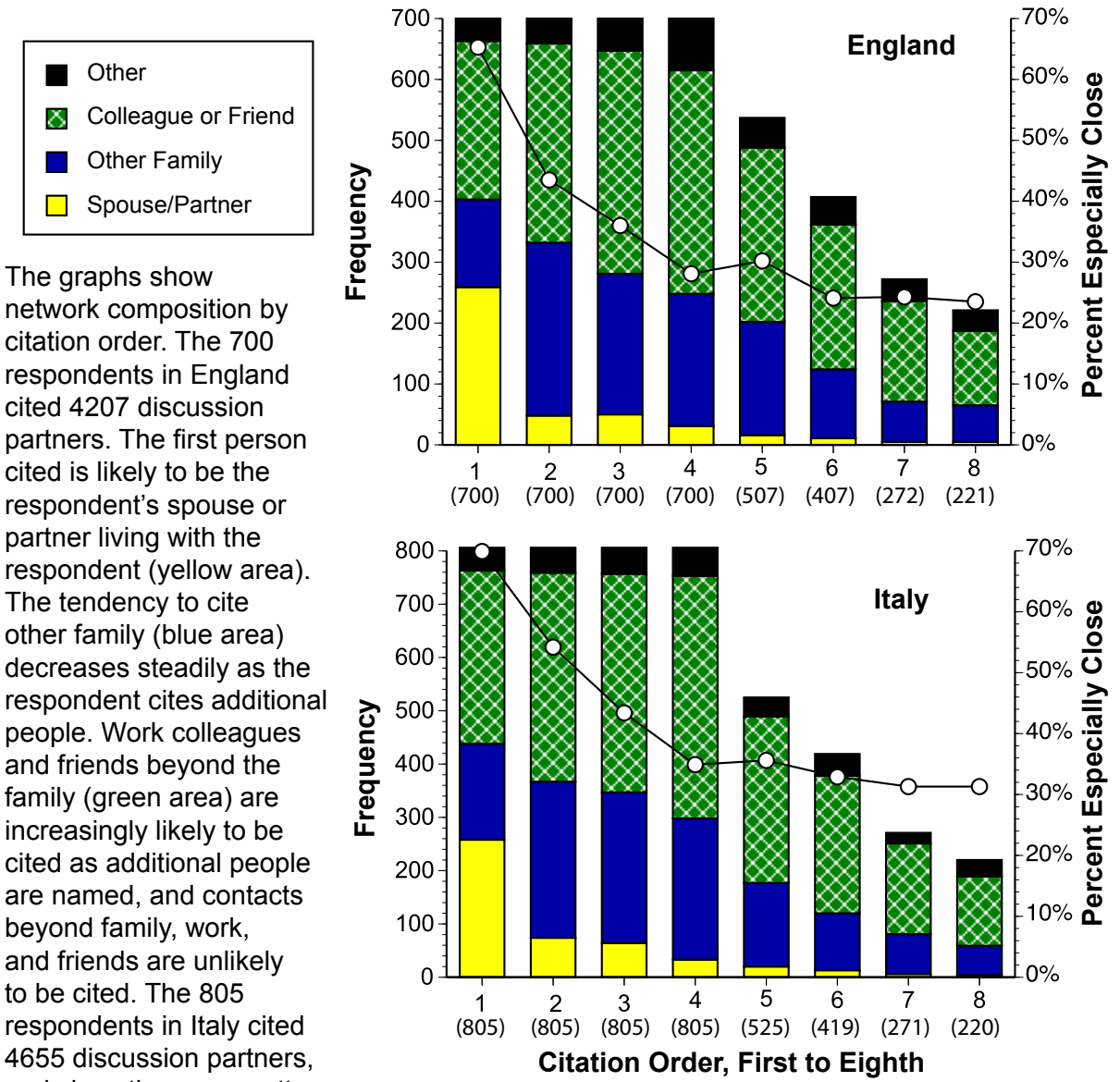
Figure 3 on the next page shows that many SNSD respondents in England and Italy named more than the minimum four discussion partners, and the probability of family being cited is concentrated in the first few citations, decreasing for people cited later. The people cited last are most likely to be work colleagues or friends beyond work. Emotional closeness is highest for the first person cited, on average, then decreases quickly through the persons cited second through fourth. People cited fourth through eighth are about equally close to the respondent.

In sum, the two graphs in Figure 3 show pattern consistent with one another, and consistent with the pattern in the GSS data gathered in face-to-face interviews. The resulting SNSD discussion networks are 59.7% contacts beyond the family in England and 59.5% in Italy (versus 45% in the GSS network data, Marsden, 1987:125). And the SNSD generator provides more data than the GSS generator on relations within the family (averages of 2.42 and 2.34 family members respectively in England and Italy, versus 1.68 in the GSS data).

Predicting Quitters. The second request for informed consent occurred after respondents completed the demographic questions on Screen 3. Therefore, to distinguish kinds of people likely to quit at the second consent, demographic attributes of people who quit at the second consent form (quitters) can be compared to the attributes of people who completed the interview (finishers). There are no significant differences in age or education between quitters and finishers. They are also similar on the odds of having a passport from another country, assessing their career progress, and acknowledging career help. They differ in marital status (3.85 logit test statistic, $P < .001$) and employment (4.31 logit test statistic, $P < .001$). In England, 34.5% of quitters are single versus 24.4% of finishers. In Italy, 44.9% of quitters are single versus 28.5% of finishers. In England, 43.3% of quitters are unemployed versus 27.1% of finishers. In Italy, 15.5% of quitters are unemployed versus 4.7% of finishers. The odds of quitting are particularly different between the countries: 5.7% of Italian entrants to the interview quit at the second consent versus 30.7% of the English entrants — which might be due to the higher unemployment among English entrants (22.2% versus 9.8% among the Italian entrants).

Network Name Interpreters

After names are generated, a series of name interpreter questions ask for: gender of each cited discussion partner (Screen 8), role relations with each person (Screen 9), energy felt from interacting with each person (Screen 10), emotional closeness to each person (Screen 11), respondent trust in each person (Screen 12), strength of connections between each pair of cited people (Screen 13), and — later in the



The graphs show network composition by citation order. The 700 respondents in England cited 4207 discussion partners. The first person cited is likely to be the respondent's spouse or partner living with the respondent (yellow area). The tendency to cite other family (blue area) decreases steadily as the respondent cites additional people. Work colleagues and friends beyond the family (green area) are increasingly likely to be cited as additional people are named, and contacts beyond family, work, and friends are unlikely to be cited. The 805 respondents in Italy cited 4655 discussion partners, and show the same pattern of network composition as additional people are named. The name generator screen would not move forward until four people had been named, so every respondent named a first through a fourth discussion partner.

The vertical axis to the right in each graph shows the percentage of cited people to whom the respondent feels "especially close" (screen 11). The decreasing lines in the graphs show that respondents begin naming their closest discussion partners, than feel equally less close, on average, to the people cited fourth through eighth.

Figure 3. Network Composition and Citation Order

interview — respondent opinion on how he or she would behave in a Trust Game with each person (Screen 19).

Gender is asked first because it is easy to answer, which facilitates transition to more difficult interpreters, and knowing contact gender allows for a measure gender homophily in a respondent's network (one indicator of a closed network).

Screen 9 elicits data on the nature of the respondent's relationship with each contact. Nuclear and extended family are distinguished, as are neighbors and colleagues

at work. There is also an “other” category to accommodate professional services or other specific connections. Each contact can be any mix of multiple things to the respondent: family and a neighbor, neighbor and a colleague, family, neighbor, and other, et cetera. These data allow us to measure network composition, such as the proportion of a respondent’s core network that is family (nuclear or extended).

The energy name interpreter in Screen 10 is a new idea. The response categories of “more energy,” “neutral,” and “less energy” from interacting with each cited person are taken from Rob Cross’ work with colleagues on emotional energy (e.g., Cross, Baker, and Parker, 2003). However, unlike Cross’ work, in which energy is used as a name generator, the concept is captured here as a name interpreter so that emotional energy can be analyzed as a correlate of network structure and trust (see discussion in Burt, Opper, and Soda, 2023).

Screens 11 and 13 elicit data on the strength of relations with and among the respondent’s cited discussion partners. These data make it possible to measure the relative strength of connections to a respondent’s family versus contacts beyond the family, and the usual measures of closure in the network around a respondent (density, constraint, betweenness).

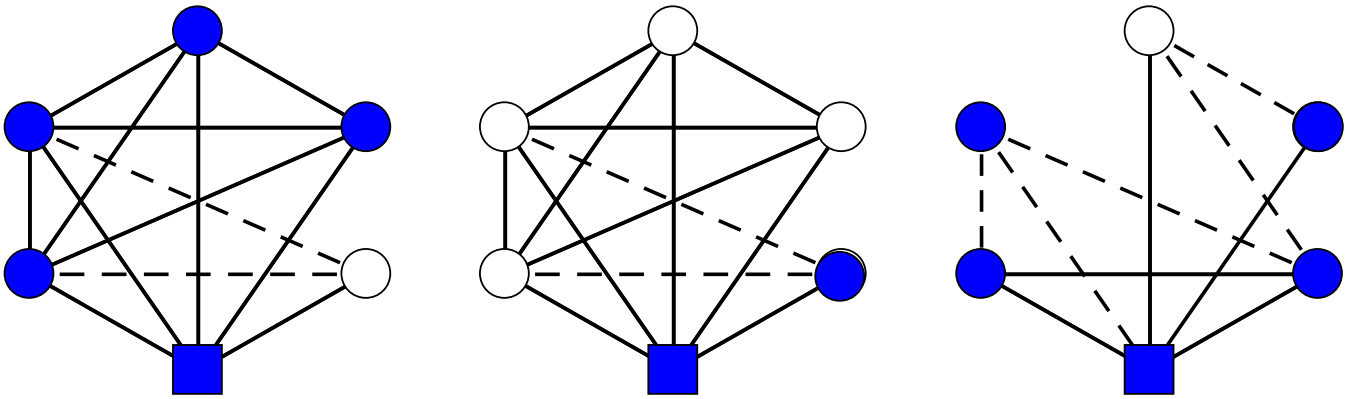
Screens 12 and 19 concern respondent trust in each contact. The question in Screen 12 asks how the respondent feels about trusting each contact, using response options from the GSS trust question in Screen 14. This question was added in the Italy survey. After playing a round of the Trust Game in Screen 17, the respondent is asked how he or she would play the Trust Game with each of the cited discussion partners. The goal is to compare trust attitude toward each cited discussion partner (Screen 12) against an attitude informed by a behavioral definition of what trust means (Screen 19).

Illustrative Core Discussion Networks

The network items are sufficient to generate images like the ones in Figure 4 on the next page. Symbols indicate people: The respondent is a square. Shading indicates family. Lines indicate relationships, thick for especially close, dashed for close, blank for less than close. Network density is computed as the average strength of connection between a respondent’s contacts (1 for especially close, .5 for close, 0 for less than close). Following Marsden (1987), I will discuss images such as the above as a respondent’s core discussion network.

The network data are less rich than the network data often collected in case studies of communities or management populations (e.g., Wellman, 1979; Fischer, 1982; Burt, 1992; Iorio, 2022), but there are mixtures of strong and weak relations possible in the data, along with variation in composition. Figure 4 is an illustrative display of the network data as they are relevant to the debate over trust eroded by strong family ties versus network closure (Burt, 2021, for review of argument and evidence). For the purposes here, reference the two arguments as family versus closure.

The two arguments agree in predicting that the respondent in network 2A will have low trust in strangers. Distrust is expected from the family argument because the



A. Dense Family
(.70 density)

B. Dense NonKin
(.70 density)

C. Sparse Family
(.30 density)

Figure 4. Illustrative Networks

respondent is deeply embedded in her family (80% of cited contacts are family and family members are connected by strong relations). At the same time, the respondent lives in such a closed network that she is also expected by the closure argument to distrust strangers, regardless of family (.70 density).

The two arguments contradict one another in network 2B. Structure is the same as in network 2A, so again the respondent is expected by the closure argument to distrust strangers. But this respondent's network is composed primarily of contacts beyond the family. There is only one family member cited, most likely the respondent's spouse, so the respondent is expected by the family argument to be open to trusting strangers.

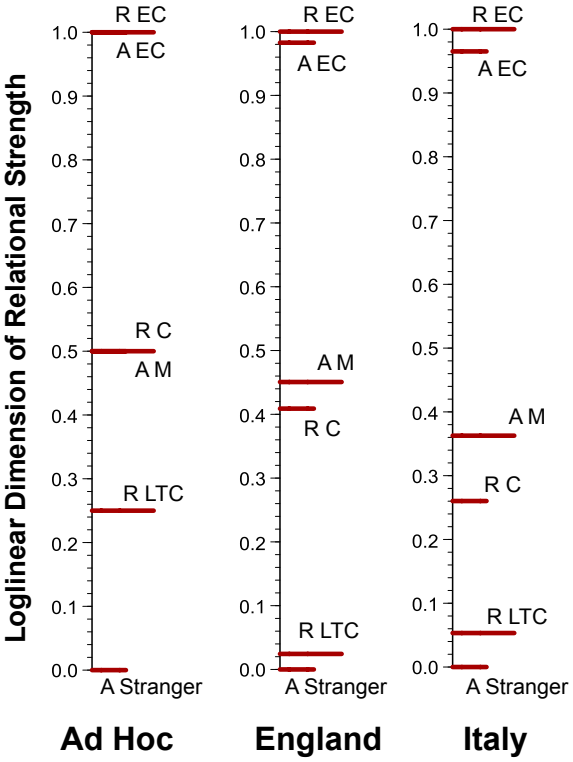
The two arguments also contradict one another in network 2C. The respondent relies on family for discussing important matters as in network 2A (network is 80% family), so she is expected by the family argument to distrust strangers. However, the family members are not especially close with one another. There are even three instances in which family members are less than close with each other. Overall, density in network 2C is much lower than it is in the other two networks (respectively .30 versus .70). By the closure argument, therefore, the respondent in network 2C is expected to be open to trusting strangers.

Scaling the Network Data

Finally, scaling is a sixth point to note about the network data. The graphs in Figure 5 on the next page show alternative scalings. "R" indicates a relation with the respondent. "A" indicates a relation between alters cited by the respondent. The first vertical axis shows an ad hoc scaling often used to compute network indices. Especially close (EC) relations are given a maximum strength of 1.0 and relations between strangers (S) are given a minimum strength of 0.0. Middling relations with the respondent (R C) or between cited discussion partners (A M) are given a strength of 0.5, and less than close relations with the respondent (R LTC) are half the strength of a middling relationship.

Balance theory can be used to scale categories by the way respondents use the categories (e.g., Burt, 2010:290-293). Details are in a technical note, but the results are given in Figure 5. The point is that respondents see a disproportionate gap between especially close relations and others. Less than close relations are almost equivalent to relations between strangers. Preliminary tests do not reveal consistently larger network effects on emotional energy when indices are computed with scaled categories, so the default is to prefer the ad hoc scaling for comparability with other research. But the sharp distinction respondents make between especially close and other relations could matter in predicting phenomena other than emotional energy.

Figure 5. Response Scaling



GSS Trust Item

As a baseline, Screen 14 contains the trust attitude question that has been used for decades in the General Social Survey (GSS) for decades. Over the years, 37% of respondents have said they trust most people (Smith et al., 2019: 394).

The item appears in numerous trust studies using the GSS as a reference point; prominently, the European Values Survey (EVS), and the World Values Survey (WVS). For example, Herreros (2015) uses World Values Survey data from 57,675 respondents in 44 countries to test the family argument discussed with respect to Figure 4 above. Trust is measured by the GSS trust item. The extent to which a respondent is deeply embedded in her family network is measured by the respondent's number of children, and a four-point opinion on the importance of family (brackets inserted, "For each of the following, indicate how important it is in your life. Would you say [family] is: Very important; rather important; not very important; not at all important?"). Number of children washes out of the analysis, but opinion on importance of family has the expected strong negative association with expressed trust in people (Herreros, 2015:348).

The question appears in Table 2 on the next page, with marginals from the 2019 GSS in the US and the EVS in Great Britain (in 2018) and Italy (in 2019), followed by marginals from the two SNSD surveys in England and Italy. Characteristic response patterns are people aggregating at the extremes (few people say "don't know") and people less likely to trust than distrust (a quarter to a third say "Most people can be trusted"). That pattern is consistent in Table 2, except for the 2023 results from England, which could be a sampling issue, though the EVS representative sample

Table 2.
Beyond your network, would you say that people can be trusted or that you can't be to careful in dealing with people?

	GSS 2018	Great Britain (EVS 2018)	Italy (EVS 2019)	England (2023)	Italy (2023)
Most people can be trusted	33.5	40.2	26.6	47.6	27.7
Don't Know	4.7	0.5	1.6	6.3	5.7
Need to be very careful	61.7	59.3	71.3	46.1	66.6
Sample N	1466	1794	2282	700	805

NOTE – Numbers are percent respondents in each category summing to the sample N. Data for the General Social Survey and European Values Surveys are available online (https://gss.norc.org and https://europeanvaluesstudy.eu).

from Great Britain is also more trusting than the Italy and GSS representative samples.

The primary difficulty with the GSS trust item is validity. What does it measure? For one thing, it is an attitude toward people in general, not individuals. Having trust in all people, ignoring the individual who is to be trusted, is certainly an indicator of optimism, perhaps too an indicator of naiveté. Second, the item elicits an attitude, not a behavior. Saying one trusts people is less risky than actually trusting people. Third, responses on the item have inconsistent association with trust behavior (Alós-Ferrer and Farolif (2019:7-9). Glaeser et al. (2000:840) conclude that game behavior from Harvard college students in introductory economics: "is at best weakly measured by typical attitudinal questions about trust, including the widely studied GSS trust question." More broadly representative of adult populations, Ermisch and Gambetta (2010:370n) "find that there is no correlation between trust behaviour as measured by our experiment and answers to survey questions about whether most people can be trusted."

Nevertheless, the GSS trust item has three virtues: face validity (it asks about trust), low cost (a respondent can answer clearly and quickly), and a large installed base of results from prior surveys. The item is included in the SNSD instrument as a low-cost baseline variable to link SNSD results to past research.

Behavioral Games

Game behavior is attractive for its face validity relative to attitude questions such as the GSS trust item. The respondent takes a concrete action that indicates trust in, or cooperation with, another person. The goal of the SNSD surveys was to experiment with network data predicting online game behavior. A one-shot Prisoner's Dilemma Game is used to measure cooperative behavior (Screen 16), and a one-shot Trust Game, adopted from Ermisch and Gambetta (2010), is used to measure trust and trustworthy behavior (Screens 17 and 19).

Screen 15 is the transition to game play. The purpose of the transition screen is to tell the respondent five things:

(1) There are three games coming up in which the respondent can win money depending on how the games are played.

(2) The respondent begins with a gift of 200 Kantar points that is hers to keep for her participation to this point. (Kantar respondents are compensated in Kantar points. One hundred points equal a euro.)

(3) In addition to the gift, the respondent will receive the highest amount she earns in any one of the games, which can be as much as another 600 Kantar points.

(4) The other person in each game is someone drawn at random from people who were interviewed earlier.

(5) Game play is anonymous — neither the other person nor the respondent has any information on one another beyond the fact that "each is an adult citizen interviewed for this research."

Two notes about point (3) on compensation: First, every respondent finishes the three games with positive earnings, regardless of game play. That fact will not be apparent during game play, but it will be apparent before the interview ends. Second, respondent earnings are defined by the game in which the respondent is most successful. This is a compromise with cost. To save costs, a colleague suggested using a lottery in which a certain number of respondents are actually compensated. A lottery would greatly lower cost, but it feels uncomfortable because privacy rules mean that the respondent cannot know for certain about compensation to other respondents.[1] In the interest of credible transparency to respondents, a summary table is displayed after all three games have been played (Screen 21). The table contains each respondent move, each corresponding move by the other person, and the resulting respondent reward — highlighting the respondent's maximum in any one game, which is credited to his or her account with the marketing research company, Kantar. Game earnings are paid at a survey's conclusion, typically a couple weeks after the survey was launched.

Point (4) on the above list is critical to research design. Before the first respondent is interviewed, about 50 pilot interviews are conducted. Respondents in the pilot interviews, distinguish them as informants, play against a computer-simulated other player. Informants serve as a pool of other players during the actual survey. For each game in each respondent interview, an informant is selected at random and that informant's move on that game is used as the other player's move for the respondent.

[1]Johnson and Mislin (2011:875) control for 10 game variables in their meta-analysis of 161 studies using the Trust Game. They find the level of trust is independent of most game variations, with only one variable rejecting the null hypothesis beyond a .001 level of confidence (-3.60 test statistic for "random payment" method in model (2)): Subjects playing under the lottery condition trust less (i.e., send significantly less money to the other player).

The informant pool of other players does three things: First, it means there is no deception when respondents are told in each game the other person is someone selected at random from earlier interviews. Game play could be simulated by having an algorithm draw virtual other players at random from a population distribution of game play. Instead, game play here is with a real person selected at random from earlier interviews.

Second, drawing the other player from the already-interviewed informant pool means that interview game play is similar to synchronous game play, which is familiar to almost everyone. When respondents play against one another in research combining a survey with behavioral games, one respondent's interview typically precedes the other's. In other words, game play is asynchronous, potentially with a substantial time gap. Combining a variation on the Trust Game with a national probability survey in Germany, Fehr et al. (2003:5-9) had respondents write moves on a sheet of paper, put the sheet in an envelope, and return the sealed envelope to the interviewer. After the survey, envelopes were matched to define which respondent played with whom, and each respondent received his or her resulting game earnings by mail. In their combination of the Trust Game with a national probability survey in Britain, Ermisch and Gambetta (2010:367) told respondents to expect to wait four weeks before they learned the other player's move. In the SNSD interview, the three other players have already made their moves, so the respondent knows all three game outcomes before the interview is concluded.

Third, using informants as a pool of other players means the data collection does not require coordination between respondent interviews (on difficulties, see Fehr et al., 2003:4-6). This is a significant advantage for survey organizations accustomed to the traditional survey design of independent respondents in area probability samples.

Cooperative Behavior - Prisoner's Dilemma Game (PDG)

Behavior in a one-shot Prisoner's Dilemma Game is used to measure respondent cooperation. Poundstone (1992: Chp. 6) describes the game's origins and its rise to wide use as a platform for trust research. Cooperation and trust are confounded in repeated play of the game, but behavior in one-shot play of the game merely indicates high or low willingness to cooperate — which of course can set a stage for later trust between the players (Kreps et al., 1995).

Following Berg, Dickhaut, and McCabe (1995), decisions are discussed in terms of investment and returns. The behavioral datum is whether the respondent chooses to INVEST or EXIT as explained on Screen 16. The cooperative move is to INVEST.

The relative game earnings are appropriate. One earns nothing without investing, and loses money if cooperation is with people who use your money instead of their own. The game is to avoid being the only investor, and hope to operate on the other person's investment. If the respondent and the other player both EXIT the game, neither one earns anything. If both INVEST, each earns 200 Kantar points. If one INVESTs and the other EXITs, the person who EXITs earns 300 Kantar points and the person who INVESTs loses 100. Outcomes meet the PDG requirement for

relative earnings from the four combinations of game moves (EI > II > EE > IE, and II > [EI+IE]/2, where the first letter is the respondent's move).

The game payoffs are a bit aggressive in that the respondent can earn nothing in the game, or even lose money. A respondent who loses money can feel like a "loser," like she has been played for a "fool," even "betrayed," the last a game-play sentiment reported in several countries among select students at prominent universities (Bohnet et al., 2008). Aversion to feeling like a loser, a fool, or betrayed, could inhibit cooperative game play. This image of game play inhibited by an aversion to felt loss in self-esteem or status, separate from monetary cost, is argued by Bohnet and Zeckhauser (2004), harkening back to Crozier (1964) on bureaucracy-prone French people not particularly liking bureaucracy, but preferring subservience to bureaucratic rules over subservience to another person, or manager self-enhancement as a driving force in organizational behavior (Pfeffer & Fong, 2005; more colloquially, think of male reluctance to ask for directions).

Sally (1995:65, 74) reports from his meta-analysis of PDG results that cooperative words in the instructions (e.g., subjects are directed to work together) significantly increase the odds of cooperation, and competitive words in the instructions (e.g., subjects are directed to do better than others) significantly decrease the odds of cooperation. From her meta-analysis of one-shot PDG results, Mengel (2017:3193-3194) reports that more likely cooperation when playing against a non-cooperative other person best predicts the lack of cooperation in a study. The game text on Screen 16 accordingly avoids competitive language. There are no "losers." The two players are not "opponents." The earnings are not "winnings." The EXIT move is not termed a "defection," as it usually is in academic discussions of the PDG.

For two reasons, cooperation seems not to be discouraged by the SNSD language. First, Table 3 on the next page shows high levels of cooperation in England and Italy. "High" is relative, of course. There is little empirical evidence directly comparable to Table 3, but there is some. I know of no study before Opper's 2012 survey of Chinese CEOs in which the Prisoner's Dilemma Game is combined with a population survey of randomly sampled respondents. Opper's respondents in 2012 cooperated in 49% of their games (Burt et al., 2022:508; and 61% in a 2021 replication survey). Those are levels well below the 82.6% and 84.7% reported in Table 3 for the SNSD surveys. Cooperation at the 80 percent level as in Table 3 is rare in one-shot Prisoner's Dilemma Games between students (37% based on average in Mengel's, 2017:3192, meta-analysis of results from lab studies of one-shot PDG), though not uncommon after two people play multiple rounds with one another (e.g., Lave, 1962:431,1965:32), or players are allowed to communicate with each other (e.g., Dawes, 1980:182 ff.).

Second, the England SNSD survey included a test for a more positive game text. Refer to the game text in Screen 16 as FORM A, and the alternative text as FORM B. FORM B was the same as FORM A except the payoffs were all positive: 200, 80, 40, 320 respectively for the four payoffs in Screen 16 (200, 0, -100, 300). Four out of five respondents were administered the FORM A game text and were cooperative in 84.4% of their Prisoner's Dilemma Games. The other one in five respondents

Table 3.
Percent Cooperative, Trusting, and Trustworthy

	England 2023	Italy 2023
1. Prisoner's Dilemma	82.6	84.7
2. Trust Game, Trustor	71.3	70.3
3. Trust Game, Trustee	61.6	63.7
Sample N	700	805

NOTE — Numbers are percent respondents in each category summing to sample N.

were administered the more-positive FORM B game text and were slightly less cooperative (75.7%, 5.96 chi-square, 1 d.f., P ~ .02). This raises a question about the game behavior data (more below when respondent time spent in each question is reported), but strengthens confidence in concluding that the game text in Screen 16 is not too negative.

Trust Behavior - The Role of Trustor in the Trust Game

Widely used in research, the Trust Game is between two people, a trustor and a trustee (see Alós-Ferrer and Farolfi, 2019, for review). SNSD respondents play both roles in two separate games. The trustor's task is defined in Screen 17. The respondent has to choose whether to INVEST her 200 Kantar points with the other player, or KEEP them. (Remember the respondent was given 200 points as a gift before she played Prisoner's Dilemma. She does not yet know how the other person behaved in the Prisoner's Dilemma, so she still has her initial 200 points.)

If the respondent chooses to KEEP her 200 points, this game ends, and she moves to the third game (Screen 20).

Choosing to INVEST is the trust decision. If the respondent chooses to INVEST, the 200-point investment is doubled to 400 points and given to the other player. The other player now has 600 points (when the 400 points are combined with the other person's initial 200). The other person now has to choose between RETURN or KEEP. If the other person also chooses RETURN, respondent and the other person split the money. Each earns 200 points in addition to their initial 200. If the other person chooses KEEP, the respondent loses her 200-point investment and the other person earns 400 points plus her initial 200. In short, the respondent has to decide whether to trust the other person to RETURN if the respondent INVESTs.

Four features make INVEST a choice of trust rather than coordination (Coleman, 1990:97-99): (1) The respondent's decision to invest gives complete control of the respondent's investment to alter to use as alter sees fit. (2) If alter is trustworthy

(shares the gains), the respondent will be better off than if he or she did not invest. (3) Respondent's decision to trust is voluntary. (4) The respondent makes his or her decision before knowing what alter will do.

The SNSD instrument uses a binary version of the usual trust game in which the respondent INVESTs some or all of her initial assets expecting the other person to RETURN some portion of the earnings (Berg et al., 1995, and 161 replication studies meta-analyzed by Johnson & Mislin, 2004). Following Ermisch and Gambetta's (2010:368-369) use of the binary trust game in their survey work, token sharing that indicates tentative trust/trustworthiness is put aside in the SNSD instrument. Respondents do not get to "risk a little" of their funds in trusting the other person, or be "a little trustworthy" in response to the other person. They either trust or they do not. They are either trustworthy, or they are not.

One might suspect that forcing an all-or-nothing decision on respondents would inhibit respondent trust, but that seems not to be a concern. Introducing the binary choice game in their British survey, Ermisch and Gambetta (2010:370) report that 43% of respondents in the role of trustor chose to SEND their initial gift to the other person in anticipation of a return (45% with a binary decision in Snijders & Keren, 1999:368). Trust in the SNSD binary Trust Game is higher still. Table 3 shows 71.3% in the England survey and 70.3% in the Italy survey.

While the decision to INVEST or KEEP is fresh in the respondent's mind, Screen 18 is a probe about the decision. This is akin to Flood's original Prisoner's Dilemma experiment in which Flood asked players what they were thinking when they made each move[2], or Ermisch and Gambetta's (2010:370) sheet of questions asking each respondent how she made her decision. In Screen 18 the respondent is asked what would be the one bit of information he or she would like to have had about the other player to guide the trust decision (e.g., race, gender, age, residence, occupation, etc.), followed by a question asking why the respondent would want that particular bit of information.

Trustworthy Behavior - The Role of Trustee in the Trust Game

In Screen 20, the respondent plays the other role in a Trust Game, the role of trustee. The trustee's task is defined in Screen 17. The respondent has been paired with a person other than the one with whom she played the role of trustor, so there is no concern about retribution from the other person. In Screen 20, the respondent is paired with someone who chose to INVEST her initial 200 Kantar points with the respondent. The respondent now has 600 points, and is asked to decide whether to RETURN half back to the other person, or KEEP it all for herself. RETURNing half the money is the choice indicating trustworthiness.

Levels of trustworthiness in the SNSD surveys are high but less out of line with prior work. Table 3 shows 61.6% trustworthy in England and 63.7% in Italy. The most

[2]Flood's original RAND report ("Some experimental games") is unpublished. Flood (1958) is an abridged version that does not contain the explanation texts. Poundstone (1992:108-116) reproduces some of the diary texts from the original report.

comparable study is the Ermisch and Gambetta (2010:370) survey, from which they report 43% trustworthy.[3]

There is an implicit contract in the trust game. The trustor invests 200 points with the expectation that the trustee will return half of the resulting 400-point gain. But there is no explicit contract, so the trustor is taking a risk assuming that the other person will honor the implicit contract to share. The trustee could justify keeping all 600 points by saying there have been no explicit statements about sharing.

An alternative game text was tested in the SNSD England survey. Refer to the game text in Screens 17 and 20 as FORM A. One in five respondents were administered a FORM B in which expectations are more explicit.

To reassure the respondent about trusting the other person (alter), the pre-choice sentence on Screen 17 — “What is your choice, INVEST or KEEP?” — was expanded in FORM B to: “The other person will be told that you sent the money trusting that the earnings would be split with you. Your task is to decide whether the other person will honor the understanding. What is your choice, INVEST or KEEP?” Table 4 at the top of the next page shows that the tendency for respondents to trust is independent of the text difference.

To encourage the respondent to be trustworthy, the pre-choice sentence in Screen 20 — “The other person will know nothing about you other than you are an adult participating in this research. How do you want to go, SHARE or KEEP?” — was expanded in FORM B to: “The other person sent the money to you trusting that you would split the increased funds, but the current decision is yours and is anonymous. The other person will know nothing about you other than you are an adult participating in this research. How do you want to go, SHARE or KEEP?” The results in Table 4 show that this text change increased the odds of a respondent sharing his or her game earnings. When explanation was limited to how the game worked (Form

[3]The SNSD trustworthy results are more consistent with slightly less comparable results in several countries for the OECD “Trustlab” project (Murtin et al., 2018). The project is related to the online survey in combining computer-implemented behavioral games with national population surveys conducted through the marketing research company, Kantar. The project is not comparable in that the project uses the Berg et al. (1995) Trust Game in which respondents decide both whether to trust as well as how much to trust. In Berg et al., ego sends some amount — $0 up to $10 — to alter (versus the binary “keep” versus “invest” choice here). The amount sent gets tripled and given to alter (who now has $10 + 3 times whatever ego sent, an upper limit of $40). Alter then decides how much to return to ego (indicating trustworthiness; again a binary “keep” versus “share” choice here). Running college students through the game, Berg et al. (1995:131) report that 94% of subjects send some amount to alter, and 16% sent the maximum of $10. On the other side of the game, 43% of alters who received one dollar or more could be termed untrustworthy in that they return to ego nothing or the one dollar. That would leave 57% as trustworthy (not too different from the 61.6% and 63.7% in Table 2 for England and Italy; and 69% of the trustworthy returned more than was sent to them). A similar design in Trustlab yields similar results. In Italy, an average of 6.00 euro were sent to alter, and alter returned an average gain of 9.44 euro to ego (Aassve et al., 2021:1040). Similar Trustlab average euro sent are reported for the US and France (Murtin et al., 2018: 27).

Table 4. Game Text Effects in Trust Game.

	Game 2: Trustor		Game 3: Trustee	
	Keep	Invest	Keep	Share
Form A	160 (72%)	396 (72%)	226 (41%)	330 (59%)
Form B	41 (28%)	103 (28%)	43 (30%)	101 (70%)
Chi-Square, 1 d.f.	0.01, P ~ .94		5.62, P ~ .02	

A), 59% of respondents shared game earnings with the other person. When the respondent was reminded that the other person had put their trust in the respondent, the percentage increased to 70%.

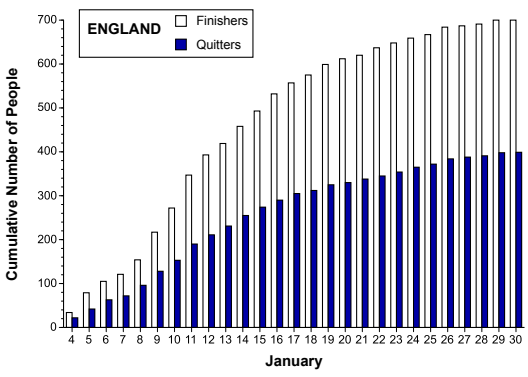
Closing the Interview

The final three screens conclude the interview. Screen 21 displays the respondent's moves in each of the three games along with moves taken by the three other people with whom the games were played, and the respondent's earnings in each game. The purpose of Screen 21 is to make the final earnings computation transparent to the respondent. Screen 22 is an opportunity for the respondent to reflect on alternative game play and to vent frustrations, if she wishes, about how she played the game. Screen 23 is a cordial good bye.

INTERVIEW

Four things stand out about how people came to, and completed, the interview. First, the smoothly decreasing numbers of people coming to the interview is consistent with people deciding independently to do the interview.[4] Second, the interview is brief for those who completed the interview in one session. The average is 9.25 minutes in England, where 89% of the one-session respondents finished within five to fifteen minutes. The average is 9.33 minutes in Italy, where 86% of the one-session respondents finished within five to fifteen minutes. Third, not all respondents finished the interview in one session. A substantial minority started the interview, then took a break before returning, on average several days later

[4]For example, the graph to the right plots cumulative number of finishers and quitters in England. The interview app was available from January 4 through 30, 2023. There is a bit of slow build-up in the initial days, suggesting a social process, but the overall pattern of continuously decreasing change implies finishers and quitters coming independently to the interview (see Coleman, Katz, and Menzel, 1957, on social versus independent adoptions.



(31% of respondents in England who completed the interview, 27% in Italy). Fourth, the specific day, or hour of the day, at which a person began the interview does not increase or decrease the odds of completing the interview, time spent in the interview, or whether the interview is stretched across multiple sessions.[5] However, beginning after office hours is an issue. Six in ten people began the interview during office hours, between the hours of 9:00 and 17:00, regardless of whether they are employed. In England, 35% of respondents who began the interview before 17:00, quit before completing the interview. That percentage jumps to 52% for those who began after 17:00 (27.33 chi-square, 1 d.f,. P < .001). Odds of quitting are lower in Italy, but the evening effect is still evident: 10% of respondents who began the interview before 17:00 quit, versus 31% of those who began after 17:00 (37.70 chi-square, 1 d.f., P < .001).

Identifying the Summary Time Variables - England

The above summary variables on interview time arose from closer study of time spent during the interview. Response in England required closer study than response in Italy, so I begin with the England data. The graph in Figure 6 at the top of the next page shows how much time was spent on each screen by the 700 respondents in England who completed the interview ("finishers" in the above discussion). Each of those respondents saw 17 screens, creating time data for a total of 11,900 respondent-screens. Time is measured on the horizontal axis in Figure 6, in five-second categories with time rounded down to the lower category. For example, the first bar in Figure 6 shows that 949 screens were viewed for less than five seconds.

Respondents moved through the screens quickly. The highest frequencies in Figure 6 are to the left of the graph. For the data plotted in Figure 6, the average time spent on a screen is 34 seconds. The median time is 25 seconds.

The time data distinguish respondents who completed the interview in one session versus respondents who left the interview, then returned later. Of the 11,900 respondent-screen times for finishers, 501 are longer than three minutes. The right hand tail of the distribution continues far beyond the graph in Figure 6. In contrast to the 34-second average for the screen times plotted in Figure 6, the 501 excluded screen times lasted an average of 181,660 seconds. That is a duration of 3,028 minutes, or 50.5 hours, or about two days. The long periods of time indicate respondents who left the interview and came back later The interview software writes a time stamp every time an interview screen is displayed. If the respondent leaves and comes back later, the screen is displayed twice, creating a time interval between viewings.

Let "one-session" respondents be the 480 who finished without spending more than three minutes on any one screen, and "multi-session" respondents be the 220 who spent more than three minutes on one or more screens. The three-minute cutoff for

[5]This sentence is based on regression models predicting a criterion variable from dummy variables distinguishing each day, or each hour of the day, at which a respondent began the interview. Logit models predict respondents who finish (versus quit) the interview, and respondents who finish the interview in one (versus multiple) sessions. OLS regression is used to predict interview duration.

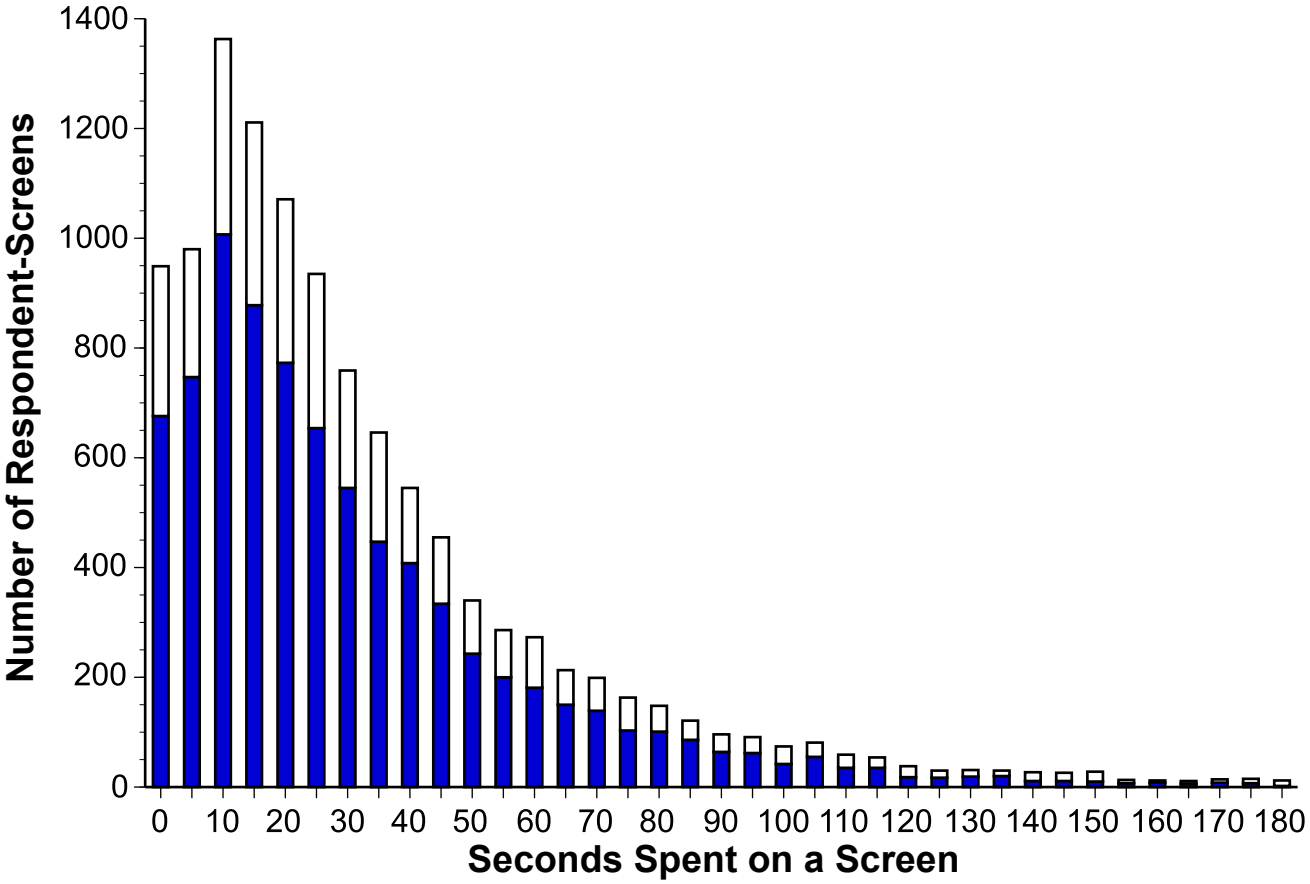


Figure 6. Seconds Spent on Each Screen by Finishers in England.
(Blue bars are times for respondents who finished the interview in one session. White bars are for those who finished in multiple sessions.)

"long viewings" is arbitrary, but reasonable in distinguishing unusual delay: Figure 6 shows few screen times beyond two minutes, and decreasing numbers as time approaches three minutes.

As a precaution, the difference between one-session and multi-session respondents is a variable to consider in data analysis, but for four reasons I am not concerned at the moment about some respondents completing the interview in multiple sessions. First, the majority of multi-session respondents only broke away once (89 of 220 spent more than three minutes on only one screen). Second, the interview software limited edits to prior responses. A respondent could go back to edit any of the background and network responses (though responses subsequent to an edited screen have to be re-entered), but game choices could not be edited. Once a respondent entered a game choice on Screens 16, 17, or 20, and pressed "Next," that game choice was no longer accessible. Third, the one- versus multi-session respondents provide data on similar size networks,[6] and the white bars in Figure 6 for multi-session respondents have a distribution similar to the solid bars for one-session respondents (multi-session respondents are consistently a few seconds slower on a screen than one-session respondents).[7]

[6]One-session respondents average 5.9 people named, versus 6.1 for multi-session respondents (1.71 test statistic for Poisson regression predicting number named, P ~ .09).

[7]Excluding the 501 screens on which multi-session resondnts spent more than three minutes, there are 11,399 screens on which one-session or multi-session respondents

Fourth, many multi-session respondents broke away at the beginning of the interview, which makes them delayed respondents more than multi-session. Rows in Table 5 on the next page distinguish interview screens listed in the Appendix. Cell entries in Table 5 show for each screen the number of people who quit at that screen,[8] the average time that one-session respondents spent on the screen, and the average time that multi-session respondents spent on the screen. The final column shows the percentage of finishers who spent more than three minutes on the row screen. Timing begins at the first display of the background demographics screen (Screen 3). The second row of Table 5 shows that of the 700 respondents who finished the interview, those who finished in one session averaged 55 seconds on Screen 3. Ninety three (13.3% of the 700) took a look at the questions and decided to do this interview later. Their average time on Screen 3 is 125,762 seconds, which means they came back at a more convenient day and a half later. Similarly, the third row of Table 5 shows that the one-session finishers averaged 79 seconds on naming discussion partners (Screen 6, on average the most time-consuming screen in the interview). One hundred and twenty of the finishers (17.1% of the 700) took a look at the name-generator screen and decided to come back later. Their average time on Screen 6 is 116,992 seconds, or about a day and a third later. The above two screens are by far the most likely to be delayed (last column in Table 5) and are delayed the longest (second to last column in Table 5). In other words, multi-session respondents were most likely to break away from the interview in the initial screens. In sum, multi-session respondents postponed completing an interview that took them slightly longer than the interviews similarly completed by one-session respondents.

Summing times across screens, the left side of Figure 7 on page 25 shows the distribution of interview times in England. Times longer than five minutes are rounded down to nearest even integer (e.g., the category labeled “16” contains durations from 16.1 to 17.9 seconds). Quitters (white bars) have a bi-modal distribution. Figure 7 shows them dropping out within the first couple minutes, but beyond the 30-minute range displayed in Figure 7, there are many who leave to go do something else then quit on their next attempt at the interview: The 238 quitters with interview times longer than 30 minutes have an average time of 8 days. Multi-session respondents (hatched areas in Figure 7) stretched the interview over periods of time when they were not in the interview. Almost half of multi-session respondents stretched the interview to beyond 30 minutes (the 100 who stretched beyond 30 minutes averaged 10 days). The best time estimate for interview duration comes from the 480 respondents who finished the interview in one continuous session (solid areas in

spent three minutes or less. These are the screens plotted in Figure 6. The one-session respondents spent an average of 30.8 seconds on each of the 11,399 screens, and the multi-session respondents spent an average of 35.9 seconds; 5.1 seconds longer. That difference is statistically significant (4.41 t-test, 4.58 test statistic for Poisson prediction). If I add dummy variables to control for average differences in time spent on different screens, then multi-session users average 7.8 seconds longer than one-session respondents (7.05 t-test, 7.53 test statistic for Poisson prediction).

[8]The disproportionate number of people quitting at the second consent in Screen 6 is apparent (310 quit, see box on page 8). After Screen 6, few people quit the interview.

Table 5. Interview Screen Times, England.

Screen/Content	N of Quitters	Finishers			
		N	Sec. One	Sec. Multi	% Multi
1-2. Splash screen & informed consent	75	700	—	—	—
3-5. Background demographics	54	700	55	125,762	13.3%
6. Expanded GSS name generator & second consent	310	700	79	116,992	17.1%
7. Name checking & editing	5	700	15	23,191	2.4%
8. Contact gender	3	700	13	22,793	2.1%
9. Role Relations with contacts	5	700	40	13,230	3.6%
10. Emotional energy from contacts	2	700	27	22,766	3.4%
11. Emotional closeness to contacts	0	700	19	8,306	2.1%
12. GSS trust in each contact (only Italy)	—	—	—	—	—
13. Closeness between contacts	4	700	64	4,323	6.6%
14-15. GSS trust & game instructions	6	700	40	21,505	8.0%
16. Prisoner's Dilemma game	2	700	32	4,783	1.3%
17. Trust Game (trustor)	2	700	26	52	1.3%
18. Reflection on Trust Game	6	700	55	14,852	4.9%
19. Trust Game with each contact	0	700	25	8,977	2.0%
20. Trust Game (trustee)	0	700	22	54	0.4%
21. Final Earnings Table	0	700	17	8,933	0.9%
22. Reflection on Game Behavior	0	700	25	8,951	1.4%
23. Thank you	0	700	0	8,911	0.7%

NOTE — In England, 1,174 people entered the SNSD interview: 474 of whom did not complete the interview (quitters) and 700 who completed the interview (finishers). Of the 700 finishers, 480 completed the interview in one session and 220 left and returned to the interview (multiple sessions, indicated by spending more than 180 seconds on one or more screens). Time is average number of seconds spent on a page, rounded down to nearest integer. The final column is the percent of screens where finishers spent more than 180 seconds (e.g., 13.3% of the 700 finishers spent more than 180 seconds on the background demographic items in Screen 3).

Table 6. Interview Screen Times, Italy.

Screen/Content	N of Quitters	Finishers			
		N	Sec. One	Sec. Multi	% Multi
1-2. Splash screen & informed consent	24	—	—	—	—
3-5. Background demographics	18	805	91	84,912	15.3%
6. Expanded GSS name generator & second consent	49	805	73	21,837	8.9%
7. Name checking & editing	0	805	13	10,183	1.7%
8. Contact gender	2	805	12	10,165	1.4%
9. Role Relations with contacts	2	805	35	10,228	3.0
10. Emotional energy from contacts	0	805	27	10,213	3.1
11. Emotional closeness to contacts	0	805	18	10,172	1.9
12. GSS trust in each contact	0	805	17	10,181	2.5
13. Closeness between contacts	7	805	57	10,257	4.7
14-15. GSS trust & game instructions	6	805	31	13,458	5.5
16. Prisoner's Dilemma game	3	805	28	60	1.1
17. Trust Game (trustor)	2	805	22	45	0.6
18. Reflection on Trust Game	6	805	55	9,620	7.3
19. Trust Game with each contact	1	805	24	5,849	2.6
20. Trust Game (trustee)	0	805	21	64	0.8
21. Final Earnings Table	0	805	16	5,774	1.0
22. Reflection on Game Behavior	0	805	18	5,781	1.1
23. Thank you	0	805	0	7,024	2.4

NOTE — In England, 925 people entered the SNSD interview: 120 of whom did not complete the interview (quitters) and 805 who completed the interview (finishers). Of the 805 finishers, 589 completed the interview in one session and 216 left and returned to the interview (multiple sessions, indicated by spending more than 180 seconds on one or more screens). Time is average number of seconds spent on a page, rounded down to nearest integer. The final column is the percent of screens where finishers spent more than 180 seconds (e.g., 15.3% of the 805 finishers spent more than 180 seconds on the background demographic screens).

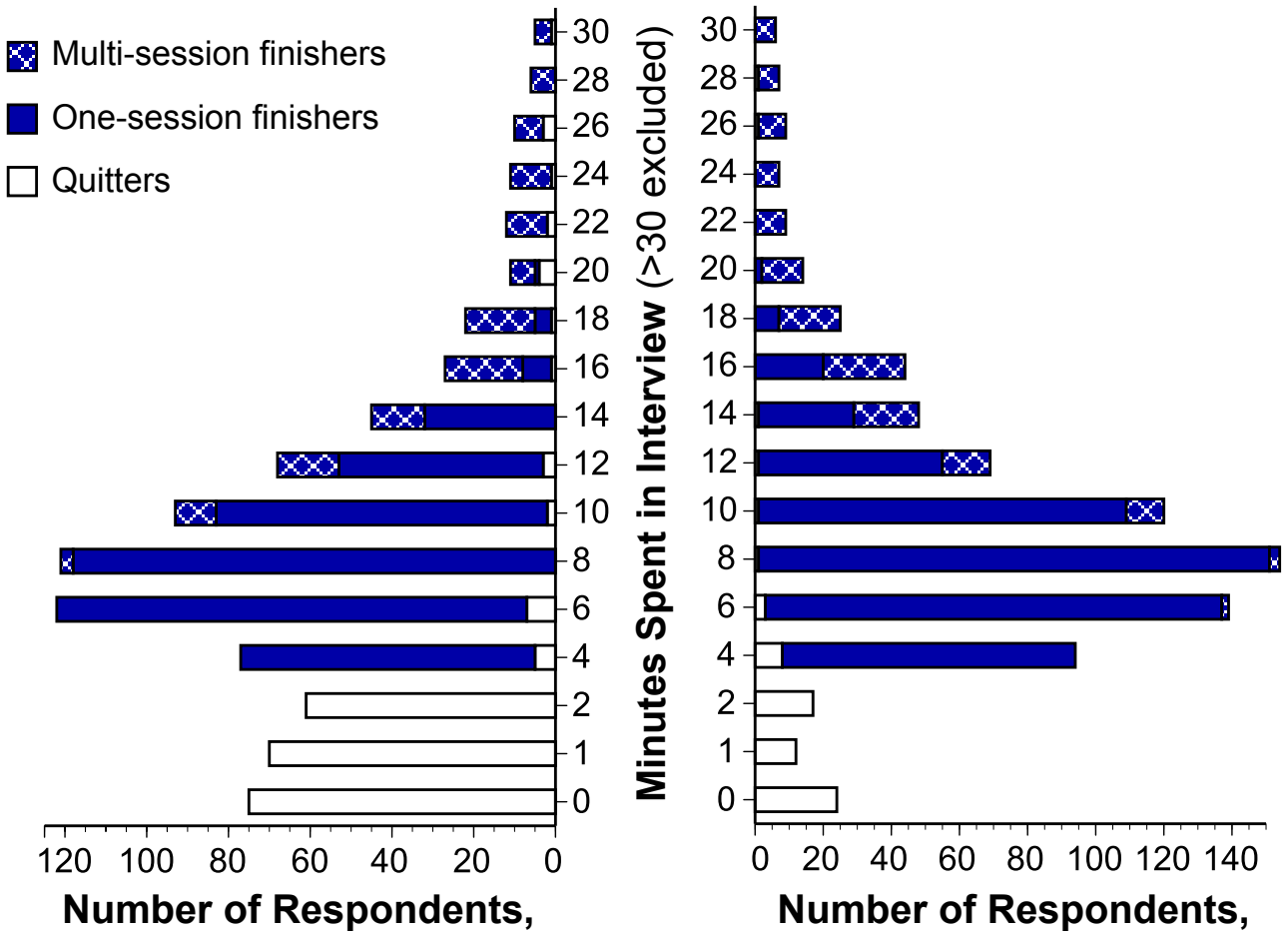


Figure 7. Interview Duration

Figure 7). They averaged 9.25 minutes to complete the interview, from a minimum of four minutes up to a maximum of 20 (median of 8.88 minutes).

Identifying the Summary Time Variables - Italy

The right side of Figure 7 shows the same graph based on interview times in Italy. The pattern is the same as in England, allowing for the relative lack of quitters in Italy. Quitters again have a bimodal distribution. The majority quit in the first couple screens (white bars at the bottom of Figure 7). A substantial minority (50 of the 120 quitters), left the interview software running for longer than 30 minutes so they do not appear in Figure 7 (average duration, 5 days). Again, multi-session respondents are individuals who spent more than three minutes on a screen. They are concentrated in the longer interviews, as in England, with the majority beyond the vertical axis (77 of 216 multi-session respondents took longer than 30 minutes; average 7 days). And again, the majority of multi-session respondents took only one break from the interview and it was when they saw the demographic background items on Screen 3. Deciding this interview will take more time than usual, so they postponed for a more convenient time. The best time estimate for interview duration in Italy comes from the 589 respondents who finished the interview in one continuous session (solid areas in Figure 7). They averaged 9.33 minutes to complete the interview, from a minimum of three minutes up to a maximum of 22 (median of 8.92 minutes).

Table 6 on page 25 provides screen time details for Italy as did Table 5 for England. As in England, the majority of people who quit the interview left at the second consent request (Screen 6, see box on page 8). The multi-session respondents were most likely to take a break from the interview when they saw the demographic background items, and their long time at that screen shows that they postponed the interview to a more convenient time.

Respondents who complete the interview in multiple sessions are engaging the interview in a manner distinct from respondents who complete it in one session, but does it matter? That is a variable to hold constant when testing for robust results.

Table 7.
Is There a Trust Tendency?

ENGLAND

	Means	GSS	Game 1	Game 2	Game 3
Means	—	.00	.84	.71	.62
GSS	.05	—	.05	.11	.17
Game 1	.79	.14	—	.20	.09
Game 2	.71	.11	.18	—	.19
Game 3	.60	07	.01	.10	—

Principal Component 33%

Principal Component 35%

ITALY

	Means	GSS	Game 1	Game 2	Game 3
Means	—	-.40	.85	.71	.63
GSS	-.35	—	.02	.04	.09
Game 1	.84	.07	—	.39	.16
Game 2	.69	.04	.12	—	.16
Game 3	.64	.05	.10	-.02	—

Principal Component 30%

Principal Component 38%

NOTE — Cells contain means and correlations for the GSS trust question (GSS, 1 for yes, 0 for don't know, -1 for no), cooperation on the Prisoner's Dilemma Game (Game 1, 1 for yes), invest when playing the trustor role in the Trust Game (Game 2, 1 for yes), share when playing the trustee role in the Trust Game (Game 3, 1 for yes). Top matrix contains results for England. Italy results are below. Results in upper diagonal of each matrix are for respondents who complete the interview in one session (480 in England, 589 in Italy). Results in the lower diagonal are for multi-session respondents (220 in England, 216 in Italy).

Here is a modest illustration: One might expect a person to display a consistent tendency toward or away from trust. Some people are cautious, even suspicious. Others are eager to cooperate, even at a hint of risk. There is a slight tendency for the one-session respondents to be more consistent in their trust responses, but it is only slight.

Table 7 shows results on constructing a single "trust" factor as a principal component of four trust variables: The GSS trust question, cooperation in the Prisoner's Dilemma Game, investing in the other person when playing the trustor role in the Trust Game, and sharing gains when playing the trustee role in the Trust Game. Responses on the GSS trust question are lower in Italy on average, but game behavior is similarly high in the two countries (as discussed with Table 3 above). Responses to the GSS question are little correlated with game behavior, as has often been reported (discussed on page 13).

What is surprising in Table 7 is weak correlation between

cooperative and trusting behavior. Altogether, a principal component index captures about a third of the variance in the four trust measures. There is more consistency across the variables for one-session respondents, but it is slight (35% versus 33% in England, 38% versus 30% in Italy). This is one of many puzzles to be solved.

Still, a third of the variation in the indicators is a large portion of available variation, and that third turns out to be a primary predictor in each measure. In the below logit regression, I pool data across games (each respondent has three records in the data file so there are 4,515 observations, 3 x 700 + 3 x 805) to predict the odds of a respondent making the cooperative/trusting game decision (logit test statistics in parentheses using Stata "cluster" option to increase standard errors for autocorrelation between game decisions by same respondent):

1.11 intercept,

.78 (9.83) cooperation in the Prisoner's Dilemma Game (game 1),

-.37 (-5.12) trustworthy behavior in the Trust Game (game 3),

.20 (4.80) respondent attitude on GSS trust item (1, 0, -1)

-.002 (-1.52) number of seconds respondent spent in this game screen,

-.07 (-0.85) respondent used multiple sessions to complete the interview,

.13 (1.67) respondent is in Italy sample, and

-.003 (-1.00) closure in network around respondent (constraint x 100).

The large effects are among the trust items. Trusting the other person in the Trust Game (game 2) is the reference category. Cooperation is higher in the Prisoner's Dilemma Game. A respondent is more likely to be cooperative in the Prisoner's Dilemma Game, which makes sense since cooperation is less risky than trust. Being trustworthy is lower in the Trust Game. A respondent is less likely to be trustworthy as trustee in the Trust Game than as trustor. That too makes sense because keeping the gains from game play is sure profit with no risk - it just isn't what a trustworthy person would do. Finally, having a positive attitude toward trusting other people (GSS trust item) is positively associated with cooperative/trust decisions.

None of the other tested predictors matter. It doesn't matter how long the respondent took on the game question. It doesn't matter whether the respondent took one or more breaks away from the interview. It doesn't matter whether the respondent is English or Italian. Closure in the network around the respondent doesn't matter.[9] In short, the cooperation/trust items have in common only a third of their variation (Table 7), but that third is statistically significant and higher than association with other predictors.

There is pattern in the association, but is more complex than the linear association estimated in the above regression. The received wisdom is to expect lower

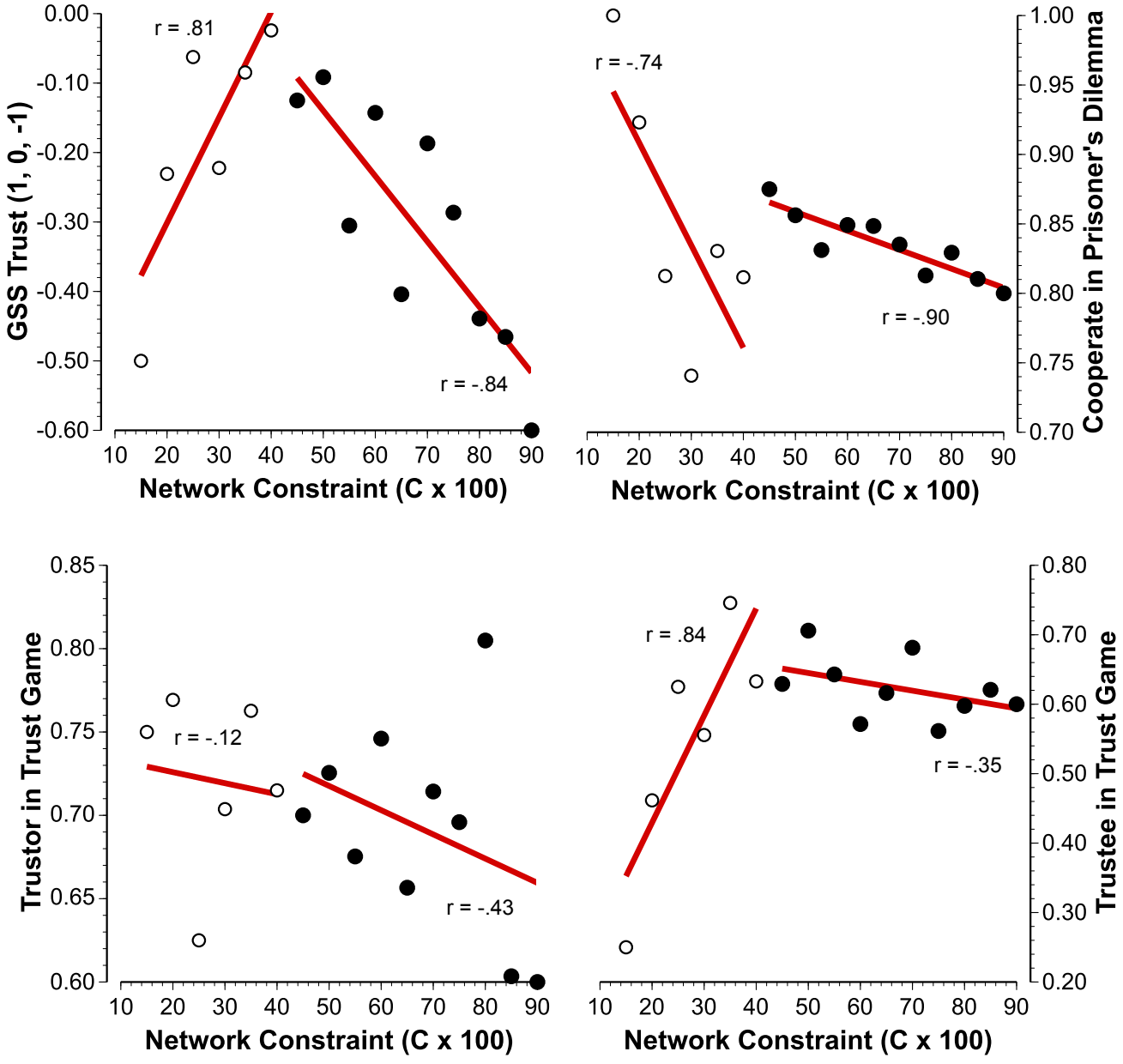
[9]The negligible association with a closed network measured by network constraint (-1.00 test statistic, P ~ .32) is similarly negligible for network size (1.15) and density (-0.38), and for constraint and density computed from scaled relations (Figure 5): -0.94 for constraint and -.06 for density.

cooperation/trust from people in closed networks (discussed earlier and in Burt, 2021). The coefficient for network constraint in the above equation is negative, but statistically negligible.

Consider Figure 8 below. The network association with cooperation/trust seems to have a threshold in the middle. At levels of constraint above 45 points, cooperation/ trust have the expected negative association with constraint (solid dots). At lower levels (hollow dots), evidence is mixed — negative association (cooperation in Prisoner's Dilemma), positive association (GSS trust item and being trustworthy), or negligible association (trusting the other player in the Trust Game). I present Figure 8 not as evidence, but as bait for further analysis. There is pattern here, but it is more complex than initially expected. Again, a puzzle to be solved.

Figure 8. Cooperation and Trust by Network Constraint

Note — Contraint is rounded to 5-point intervals.



REFERENCES

Aassve, Arnstein, Conzo, Pierluigi, & Mattioli, Francesco (2021) “Was Banfield right? New insights from a nationwide laboratory experiment.” *Journal of Regional Science* 61(5): 883-1139.

Alós-Ferrer, Carlos, & Farolfi, Federica (2019) “Trust games and beyond.” *Frontiers in Neuroscience* 13: 887.

Berg, Joyce, Dickhaut, John, & McCabe, Kevin (1995) “Trust, reciprocity, and social history.” *Games and Economic Behavior* 10(1):122-142.

Bohnet, Iris, & Zeckhauser, Richard (2004) “Trust, risk, and betrayal.” *Journal of Economic Behavior and Organization* 55: 467-484.

Bohnet, Iris, Greig, Fiona, Herrmann, Benedikt, & Zeckhauser, Richard (2008) “Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States.” *American Economic Review* 98(1): 294-310.

Burt, Ronald S. (1984) “Network items and the General Social Survey.” *Social Networks* 6: 293-339.

Burt, Ronald S. (1986) “A note on sociometric order in the General Social Survey network data.” *Social Networks* 8: 149-174.

Burt, Ronald S. (1991) “Measuring age as a structural concept.” *Social Networks* 13:1-34.

Burt, Ronald S. (1992) *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Burt, Ronald S. (2010) *Neighbor Networks: Competitive Advantage Local and Personal*. New York: Oxford University Press.

Burt, Ronald S. (2021) “Note on family versus network closure as foundation for distrust and lack of cooperation.” Research Note, Bocconi University. (http://ronaldsburt.com/research/files/FNC.pdf)

Burt, Ronald S., Opper, Sonja, & Holm, Hakan J. (2022) ”Cooperation beyond the network.” *Organization Science* 33(2): 495-517.

Burt, Ronald S., Opper, Sonja, & Soda, Giuseppe (2023) “Energy and structural holes.” Working paper presented at ION X, University of Kentucky.

Coleman, James S., Katz, Elihu, & Menzel, Herbert (1957) “The diffusion of an innovation among physicians.” *Sociometry* 20(4): 253-270.

Coleman, James S. (1990) *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Connor, Kathryn M., & Davidson, Jonathan R. T. (2003) “Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC).” *Depression and Anxiety* 18:76-82.

Cross, Rob, Baker, Wayne, & Parker, Andrew (2003) “What creates energy in organizations?” *Sloan Management Review* 44(4): 51-56.

Crozier, M. 2010 [1964] *The Bureaucratic Phenomenon*, (with a new introduction by E. Friedberg). New Brunswick, NJ: Transactions.

Dawes, Robyn M. (1980) “Social dilemmas.” *Annual Review of Psychology* 31: 169-193.

Ermisch, John, & Gambetta, Diego (2010) “Do strong family ties inhibit trust?” *Journal of Economic Behavior and Organization* 75: 365-376.

Fehr, Ernst, Fischbacher, Urs, Von Rosenbladt, Bernhard, Schupp, Jurgen, & Wagner, Gert G. (2003) “A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys.” CESifo Working Paper Series, No. 866; IZA Discussion Paper, No. 715. (https://www.iza.org/publications/dp/715/a-nation-wide-laboratory-examining-trust-and-trustworthiness-by-integrating-behavioral-experiments-into-representative-surveys)

Fischer, Claude S. (1982) *To Dwell Among Friends: Personal Networks in Town and City*. Chicago, IL: University of Chicago Press.

Fischer, Claude S. (2009) “The 2004 GSS finding of shrunken social networks: An artifact?” *American Sociological Review* 74(4): 657-669.

Flood, Merrill M. (1958) “Some experimental games.” *Management Science* 5(1): 5-26.

Glaeser, Edward L, Laibson, David I., Scheinkman, Jose A., & Soutter, Christine I. (2000) “Measuring trust.” *Quarterly Journal of Economics* 115 (3):811-846.

Herreros, Francisco (2015) “Ties that bind: Family relationships and social trust.” *Rationality and Society* 27(3): 334-357.

Iorio, Alessandro (2022) “Brokers in disguise: The joint effect of actual brokerage and socially perceived brokerage on network advantage.” Administrative Science Quarterly 67(3): 769-820.

Johnson, Noel D., & Mislin, Alexandra A. (2011) “Trust games: A meta-analysis.” *Journal of Economic Psychology* 32(5): 865-889.

Kreps, David M., Milgrom, Paul, Roberts, John, & Wilson, Robert (1982)‘‘Rational Cooperation in the Finitely Repeated Prisoner’s Dilemma.’’ *Journal of Economic Theory* 27(2): 245–252.

Lave, Lester B. (I962) “An empirical approach to the prisoners’ dilemma game.” *Quarterly Journal of Economics* 76(3):424-36.

Lave, Lester B. (I965) “Factors affecting co-operation in the prisoner’s dilemma.” *Behavioral Science* IO(1): 26-38.

Marsden, Peter V. (1987) “Core discussion networks of Americans.” *American Sociological Review* 52(1): 122-131.

Marsden, Peter V. (2011) “Survey methods for network data.” Pp. 370-388 in *The SAGE Handbook of Social Network Analysis*, edited by John P. Scott and Peter J. Carrington. Thousand Oaks, CA: Sage.

Marsden, Peter V., Fekete, Maleah, & Baum, Derick S. (2021) “Egocentric network studies within the General Social Survey: Measurement methods, substantive findings, and methodological research.” Pp. 519-551 in Mario L. Small, Brea L. Perry, Bernice A. Pescosolido, & Edward B. Smith (Eds.), *Personal Networks: Classic Readings and New Directions in Egocentric Analysis*. New York: Cambridge University Press.

Mengel, Friederike (2017) “Risk and temptation: A meta-study on prisoner’s dilemma

games.” *Economic Journal* 128(616):3182–3209.

Merluzzi, Jennifer, & Burt, Ronald S. (2013) “How many names are enough? Identifying network effects with the least set of listed contacts.” *Social Networks* 35: 331-337.

Murtin, Fabrice, et al. (2018) “Trust and its determinants: Evidence from the Trustlab experiment.” OECD Statistics Working Papers, No. 2018/02, OECD Publishing, Paris (https://doi.org/10.1787/869ef2ec-en)

Paik, Anthony, & Sanhagrin, Kenneth (2013) “Social isolation in America: An artifact.” *American Sociological Review* 78(3): 339-360.

Perry, Brea L., Pescosolido, Bernice A., & Borgatti, Stephen P. (2018) *Egocentric Network Analysis*. New York: Cambridge University Press.

Pfeffer, Jeffery, & Fong, Christina T. (2005) “Building organization theory from first principles: The self-enhancement motive and understanding power and influence.” *Organization Science* 16(4): 372-289.

Poundstone, William (1992) *Prisoner’s Dilemma*. New York: Penguin.

Sally, David (1995) “Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992.” *Rationality and Society* 7(1): 58-92.

Smith, Tom W., Davern, Michael, Freese, Jeremy, & Morgan, Stephen L. (2019) *General Social Surveys, 1972-2018: Cumulative Codebook*. Chicago, IL: NORC.

Snijders, Chris, & Keren, Gideon (1999) “Determinants of trust.” Pp. 355-385 in Budescu, D. V., Erev, I., & Zwick, R. (Eds.), *Games and human behavior: Essays in honor of Amnon Rapoport*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Vaishnavi, Sandeep, Connor, Kathryn M., & Davidson, Jonathan R. T. (2007) “An abbreviated version of the Connor-Davidson Resilience Scale (CD-RISC), the CD-RISC2: Psychometric properties and applications in Psychopharmacological trials.” *Psychiatry Research* 152 (2-3): 293-297.

Wellman, Barry (1979) “The community question: The intimate networks of East Yorkers.” *American Journal of Sociology* 84(5): 1201-1231.