

Two-route models of trust in sources of risk communication

Nigel Harvey and Matt Twyman

University College London

Acknowledgement: This research was supported by Economic and Social Research Council Grant R 000 230114.

Abstract

We discuss trust in different domains, focusing on studies of risk communication. In this context, trust appears to have two main components: trust in the motives (benevolence, integrity, honesty) of the advisor and trust in the advisor's competence. Siegrist, Earle and Gutscher's (2003) refer to these components as (social) trust and confidence, respectively. In their Trust-Confidence-Cooperation (TCC) model, these separate types of trust jointly determine an agent's degree of cooperation with the source of advice (expressing trust in it, using its advice to form personal assessments of risk, acting on the basis of its advice). We discuss evidence suggesting that different types of cooperative behaviour can be dissociated and assess the implications of this both for the TCC model and for the study of risk communication.

Trust is a focus of study in many areas of the social sciences, including economics (Camerer, 2005), politics (Markova, 2004), organizational behaviour (Kramer and Tyler, 1996) and sociology (Fukuyama, 1995). It has been defined in various ways but is broadly seen as a willingness to make oneself vulnerable to the views, decisions or actions of another person or an organization. Here we shall be concerned with its psychological foundations.

Psychological foundations of trust: Studies in three domains

The psychological basis of trust has been systematically studied in three main domains. First, behavioural game theorists have used various types of games to study the conditions that influence how much trust people are willing to place in someone else. These games are useful in that they allow the other person's trustworthiness to be measured as well. For example, in a basic trust game (Berg, Dickhaut and McCabe, 1995) between two players, one player is endowed with an amount of money by the experimenter and must decide what proportion of that amount they will hand over to the second player for investment. Whatever gets handed over is multiplied by some factor (e.g. it is doubled). The second player then decides how much of this new amount they will hand back to the first player. From the perspective of a traditional economic rationalist, the second player should hand nothing back and, realizing this, the first player should invest nothing in the first place. However, if the second player can be trusted, it is sensible to for the first player to invest. Thus, within this game, the amount of money handed over for investment can be regarded as a measure of the first player's trust in the second one. Conversely, the amount handed back after the investment has had its effect can be seen as a measure of the second player's trustworthiness.

Studies using the basic trust game and multistage versions of it in which invested money shuttles back and forth between the two players ("centipede games") have provided a rich source of data for theorists. Also, Fehr and Gächter (2000) have extended these games to model trust in different types of labour markets and their findings do appear to predict effects of different types of organizational design in real firms (Baron, Hannon and Burton, 2001).

A second domain in which the psychological determinants of trust have been studied empirically is that of electronic commerce (e.g. Grabner-Kräuter and Kaluscha, 2003; Corritore, Kracher and Wiedenbeck, 2003). Internet shoppers must decide whether they can trust those running commercial websites to provide the goods once they have been paid for (or once credit card details have been provided). If they are not trusted by consumers, trustworthy electronic business will fail. Conversely, the effects of consumers' trusting untrustworthy retailers may be detrimental for the success of electronic commerce in general. Thus, studies have investigated relations between website characteristics, consumers' ratings of trust in retailers, and the likelihood of a completed transaction (e.g. Pavlou, 2003).

The third domain in which investigators have examined the psychological determinants of trust is that of risk communication. Increasingly, we seem to expect governments to manage the various risks to which we are exposed. Such management can be accomplished in two main ways. Legislation can be enacted to control risks. For example, laws against using handheld mobile telephones while

driving have been introduced in many countries. Alternatively, levels of risk associated with various hazards can be communicated to the public via electronic and traditional media. For example, people have been informed of the likelihood of dying from smoking-related diseases in this way.

Theoretical developments in these three domains are not especially well integrated. (Studies of trust in e-commerce have cited work on trust in risk communication but they are an exception.) One reason for this may be that those working in different domains have used different approaches to measuring trust and, as a result, have studied rather different forms of trust that need to be understood in somewhat different ways. Specifically, those working within the context of behavioural game theory have focused on behavioural measures of the placement of trust – most commonly, the amount of money handed over to another player. In contrast, researchers into trust in e-commerce and risk communication have typically collected verbal statements of levels of trust from participants in their studies. It may be that quite different theories need to be developed to account for the behavioural placement of trust and for verbal statements of trust.

To examine this possibility in more detail, we shall focus on theories of trust in risk communication. We shall briefly review these theories, isolate one which is particularly well specified, and then consider the adequacy of this model in the light of both behavioural and verbal data.

Trust in sources of risk communication

People say that they have more trust in some sources of risk communication than others. Many studies have shown that people claim to have less trust in government sources of information than in consumer organizations, family, and friends. For example, Frewer, Howard, Hedderley and Shepherd (1990) asked people within a semi-structured interview to rate their distrust in 15 different information sources. Tabloid newspapers were least trusted, followed by government agencies, friends, pressure groups, quality newspapers, television news and documentaries, and consumer groups. Medical doctors and university scientists were the most trusted sources of information. However, the ranking of trust in different sources depended to some extent on the type of risk that was under consideration (e.g. alcohol use, high-fat diet, pesticide residues).

Factor analytic studies of questionnaires containing items relating to different potential influences on trust have identified different features of sources of information as factors that determine trust. For example, Renn and Levine (1991) identified competence, objectivity, fairness, faith (goodwill) and consistency, whereas Mayer, Davis and Schoorman (1995) isolated ability, benevolence and integrity. As a result of these studies, various models of trust in risk communication have been developed. Differences between these models reflect differences in the way two broad (groups of) factors determining trust have been labelled.

Thus, according to one type of two-route model, the first factor can be regarded as a competence component and the second as an affective component of trust (Mettay, 1999). In this way, the processing of trust can be related to the well established tradition of distinguishing cognitive and affective processing of information (e.g. Zajonc, 1980). Over the last few years, the importance of affective processing in risky

decision-making has been recognized in theoretical developments such as the risk-as-feelings hypothesis (Loewenstein, Weber, Hsee and Welch, 2001) and identification of the affect heuristic (Slovic, Finucane, Peters and MacGregor, 2002; Finucane, Alhakami, Slovic and Johnson, 2000). More recently, Finucane and Holup (2006) have developed a two-route (affective/cognitive) model of the processing of risk information in which differences in perceived risk can be explained in terms of processing in either route or in terms of the way the results of each type of processing are combined.

McAllister (1995) used the affect-cognition dichotomy as the basis for his dual-process model of trust. The cognitive route relies on processing information about the knowledge, competence, reliability, and dependability of the source. In contrast, the affective foundations of trust depend on emotional bonds: "People make emotional investments in trust relationships, express genuine care and concern for the welfare of partners, believe in the intrinsic virtue of such relationships, and believe that these sentiments are reciprocated" (McAllister, 1995, p. 26). Others have developed and generalized this dual-process approach (e.g. Rousseau, Sitkin, Burt and Camerer, 1998).

From our perspective, the problem with using the affect-cognition dichotomy to identify the two routes of processing that determine trust is that it is more appropriate for explaining how someone develops trust in a person with whom they interact than in an agency from which they receive risk information. This is because McAllister's (1995) model was originally developed to explain interpersonal trust between managers and professionals in organizations. Within it, affect-based trust increases with interaction frequency because frequent interactions showing behaviour that is personally chosen rather than role-prescribed and demonstrably interpersonal care and concern rather than enlightened self-interest is needed to foster emotional bonds. However, people do not usually have these types of interaction with the government agencies and consumer organizations from whom they obtain risk information.

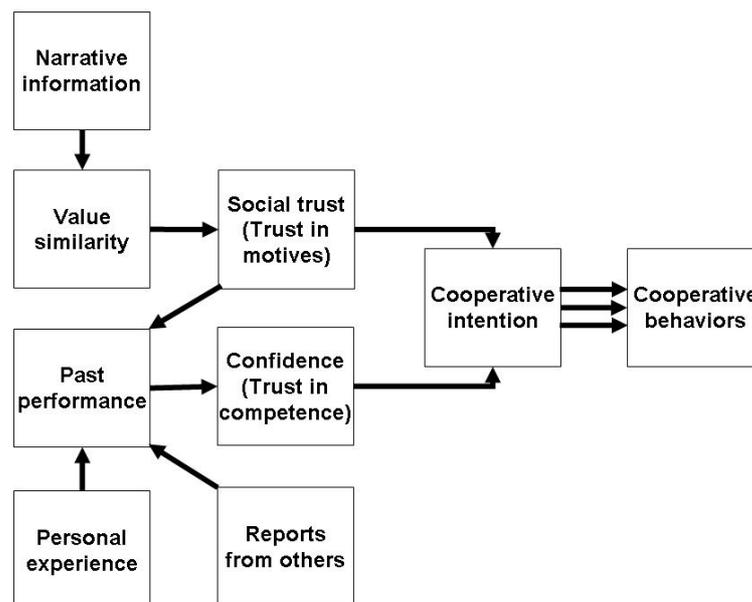
An alternative dichotomy can be derived from the results of the factor analytic studies that is more appropriate for modelling trust in risk communication. Siegrist, Gutscher and Earle (2005) argue that the factor related to the benevolence, honesty and integrity of the source corresponds to trust in motives of the source rather than to affect-based trust. Also, the factor related to ability, competence and knowledge of the source corresponds to confidence in the source rather than cognition-based trust. We can assess the motives and competence of an agency with which we are not in two-way communication. Thus Siegrist et al.'s (2005) two-route of trust is more appropriate for our purposes than the approach adopted by McAllister (1995).

It is worth emphasizing, however, that Siegrist et al.'s (2005) model does not exclude affective processing from the development and placement of trust. Trust in motives depends on an assessment by the receiver of the information of how similar the values of the source are to their own. If the source's values, intentions and goals are similar to those of the receiver of information, trust in the source's motives will be higher. Clearly, feelings of empathy could be involved in this process.

Siegrist et al.'s (2005) model is an appropriate two-route model of trust for considering issues related to risk communication. It is also well specified and has some empirical support. In the next section, we consider details of their model.

The trust-confidence-cooperation model of risk communication

The figure below shows the essential features of Siegrist et al.'s (2005) Trust-Confidence-Cooperation (TCC) model. Receivers' trust in the motives ('social trust') of a source depends on how similar they judge that source's values to be to their own. This judgment depends, in turn, on narrative information provided by the source. Their trust in the competence ('confidence') in the source depends on their assessment of the quality of information provided by the source in the past. This, in turn, depends on their personal experience of the source and/or on reports that they receive from others about the quality of information that it has provided in the past. Social trust and confidence in the source jointly determine the level of cooperation that the receiver intends to have with the source. This intention can produce a variety of cooperative behaviours, such as expressing trust in the source, accepting advice about risk levels from a source (risk perception) and acting on the basis of that advice (risk acceptance or rejection).



As the figure shows, one feature of the TCC model is that level of social trust acts to filter the performance information that determines confidence in the source. Thus if someone accepts that a source has benevolent motives they are more likely to assess poor performance by the source charitably. Siegrist et al. (2005) include this filtering feature in their model because there is evidence from studies of impression formation (De Bruin and Van Lange, 1999, 2000) that social desirability tends to dominate intellectual desirability. Because of this asymmetry, processing related to trust-in-motives can influence processing related to trust-in-competence but not vice versa.

A second important feature of the TCC model is that the two routes (trust and confidence) converge on a final common pathway. As a result, cooperative

behaviours should differ or change in a coherent way. For example, if someone expresses more trust in a consumer organization than in a government agency, then they should also be more inclined to take advice from the consumer organization than from the government agency. Similarly, if someone says that they have lost trust in a government agency, then they should be less inclined to take advice from that agency than they were before.

These two features of the TCC model are important. We shall refer to them as the asymmetric influence assumption and the final common pathway assumption, respectively. Some of our own work has been concerned with testing these assumptions. Before describing it, however, we shall outline some of the evidence that supports the TCC model.

Evidence supporting the TCC model

Earle and Cvetkovich (1999) presented people with a fictional story (framed as a simulated newspaper report) about a prospective high-level nuclear waste management policy. The story described in some detail a new decision-making procedure that would be used. (The nature of this procedure varied between conditions.) Respondents were told that the proposed new federal organization that would carry out the procedure was the Nuclear Waste Management Agency (NWMA). They were asked to rate the extent to which they would trust the NWMA. They were also asked to complete a six-item scale that asked them to assess how similar their values, goals, views, and thoughts were to those of the NWMA. A similarity-of-values index was extracted from these six items. In a first study, this index had a 0.66 correlation with the level of trust that people expressed in the NWMA. In a second study, the corresponding correlation was 0.68. Given that expressions of trust are cooperative behaviours, these results show that similarity of values influences such behaviours, just as the model suggests that it should (via the mediation of trust-in-motives).

Harvey, Twyman and Harries (2007) replicated Earle and Cvetkovich's (1999) finding for a wide variety of occupational, recreational and transport risks and with various consumer organizations and government agencies acting as risk communicators. Generally, however, the correlations between stated trust and the similarity of values index that they obtained, though significant, were rather lower than those reported by Earle and Cvetkovich (1999). Their median values were .35 (Study 1) and .36 (Study 2).

Harvey et al.'s (2007) study also found that participants judged government agencies to be more accurate advisors than consumer organizations but assessed their values to be closer to those of consumer organizations. Various different measures of trust showed either no difference between those two types of advisor or a tendency to place more trust in consumer organizations. This pattern of results can be interpreted within Siegrist et al.'s (2003) model in one of two ways. Either trust in motives (determined by similarity of values) filters the performance information that determines trust in competence in the way that they suggest (before the two routes converge in producing cooperative intentions) or conflicts in which the two types of trust indicate different types of cooperation are resolved by some sort of weighting process (when the routes converge into the final common pathway).

Siegrist, Earle and Gutscher (2003) used a questionnaire to measure the constructs in the model in the context of the electromagnetic field risks associated with mobile phone antennae. Each construct had a number of items associated with it. For example, one of those associated with confidence in the model was: "Mobile phone companies possess the competence necessary to assess any health risks associated with their antennas." Respondents rated their agreement with each item on a five-point scale. The plausibility of the TCC model was tested using structural equation modelling. The structure of the model was broadly supported. However, the hypothesized link between past performance and confidence was weak and there was a direct effect of social trust on confidence. Siegrist et al. (2003) argued that these differences between model and data occurred because respondents knew too little about the electromagnetic field risks to use past performance of the source as a guide to their confidence in it. As a result, they had to use their trust in the source's motives as a proxy for past performance.

To test this familiarity-based account, Earle and Siegrist (2006, Study 1) carried out a study similar to that of Siegrist et al. (2003) but using Arctic Oil exploration risks (unfamiliar to respondents) and risks associated with freeway expansion in Seattle (familiar to respondents). Results for Seattle traffic were similar to those obtained by Siegrist et al. (2003). However, because respondents had enough familiarity with the risks to use past performance to assess confidence, there was a stronger link between past history and confidence – just as predicted. Results for the Arctic Oil exploration scenario had been expected to be similar to those obtained by Siegrist et al. (2003) because respondents were unfamiliar with the risks in both cases. However, there was a strong link between past performance and confidence but confidence had no effect on cooperation. It is unclear why the effects of being unfamiliar with the risks should have had different effects in the Siegrist et al. (2003) and in the Earle and Siegrist (2006) studies.

Tests of the final common pathway assumption

As we pointed out above, the final common pathway assumption implies that cooperative behaviours should differ or change in a coherent way. For example, someone who expresses more trust in one source of risk information than in another should also be more inclined to take advice from the source that they say they trust more. However, O'Neill (2002) has argued that verbal expressions of trust may be dissociated from the way that trust is revealed in behaviour: "Are the opinions we divulge to pollsters backed up by the ways in which we *actively* place our trust in others... The evidence suggests that we still consistently place trust in many of the situations and professions we profess not to trust" (pp. 11–13). She argues that this is not because people have no alternative but to behave the way they do. They place their trust in agencies that they say they do not trust even though they are free to place it elsewhere. O'Neill's (2002) distinction between expressed and placed trust is not dissimilar to the distinction between that Samuelson (1948) introduced into economics between stated and revealed preferences. Indeed, adopting his terminology, we can refer to stated trust as a level of trust described in words or numbers (e.g. via a rating scale) and to revealed trust as a level of trust that is evident in (non-verbal) behaviours.

Clearly, there are a number of reasons why stated and revealed trust may dissociate. For example, consider people who shop for their food in supermarkets: their behaviour reveals that they trust supermarket food. However, they may still have serious concerns about various aspects of the food that they buy. They may be worried about the safety of genetically engineered food and suspect that supermarkets include some food of this type among their produce without clearly labelling it as such. As a result, when questioned by pollsters, they may say that they do not trust supermarket food and that the reason for this is that they are not sure what is in it. Even though their worries are not sufficient to change their supermarket shopping habits, they will express them when given the opportunity to do so by a pollster – perhaps in the knowledge that their concerns will eventually be communicated to supermarket managers, who may act on them. In other words, people may use their statements of trust as opportunities to express their concerns or as warning signals that they will withdraw their actual placement of trust if some aspect of the situation does not improve.

There are, however, other reasons why stated trust and revealed trust may dissociate. People may lack full insight into how they place trust. For example, when deciding whether to expose themselves to some hazard (motorcycling, taking heroin), they may not be aware of how they are weighting different sources of information about the risks associated with that hazard. When asked about the level of risk, they may recall a particular media report or conversation with a friend and say that is a source that they trust. However, their decision about whether to take a particular risk may well depend on a much wider sample of advice – without them necessarily being aware of this.

Cognitive psychologists have carried out experimental studies that do appear to confirm that people do not always have full insight into the processes underlying their behaviour. We may be no more aware of our mental processes than we are of our digestive ones. As a result, the 'explicit' processes underlying what we say about why we make a judgment or decision may be quite different from the 'implicit' processes that actually determine that judgment or decision. For example, Berry and Broadbent (1984, 1988) either gave people practice at a decision-making task or provided them with verbal information about how to perform it. Those with experience at the task learned to perform it but could say little about how it should be performed or what they were doing to achieve their success. Conversely, those given verbal instructions were able to say something about how the task should be performed but showed little evidence of being able to transform this theoretical knowledge into success at the task.

Most people are familiar with dissociations between implicit and explicit processing in the domain of motor skills. Telling someone how to ride a bicycle or walk a tightrope based on information about the physics of the task can provide them with correct theoretical knowledge but no competence in the task. However, after years of practice at the task, people may well be highly competent at it while having little knowledge of the physics underlying their performance: as a result, they cannot say much about what they are doing to achieve success. Work by Berry and Broadbent (1984, 1988) and others before them (Reber, 1967) is important because it suggests that we should consider the possibility that similar dissociations between implicit and explicit processing can occur in cognitive tasks.

Harvey et al. (2007) provided people with advice from a government agency and a consumer organization about risks associated with a variety of hazards and asked them to use this advice to form their own assessment of the risks. The emphasis or weight that they placed on each type of advisor was derived statistically and used as an estimate of how much they placed their trust in them. In other words, this was a behavioural measure of revealed trust. After people had made their estimates, they were asked to use a rating scale to express their level of trust in each type of advisor: consumer organizations were trusted more than government agencies: this served as a measure of stated trust.

In one experiment, revealed trust showed a difference between the two types of advisor: however, there was no significant difference in the levels of stated trust in each type of advisor. Under different conditions, a second experiment produced results in the opposite direction: there was no difference in revealed trust between the two types of advisor but stated trust was higher in consumer organizations than in government agencies. This 'double dissociation' between stated and revealed trust provides reasonably strong evidence that the cognitive processes underlying the behavioural placement of trust are distinct from those underlying the verbal expression of trust. Because it was obtained within an artificial experimental situation, it is not easy to explain it in terms of the sort of signalling of concern that we discussed above in the context of poll results.

This dissociation between stated and revealed trust implies that the final common pathway assumption of Siegrist et al.'s (2003) study may need some revision. However, this issue is not relevant to processes prior to formulation of cooperative intention. However, the assumption of asymmetric influence is relevant to those processes and it is to that that we turn next.

Tests of the asymmetric influence assumption

According to the TCC model, activity in the trust-in-motives route can influence that in the trust-in-competence route but not vice versa. Thus, high trust in an agent's motives will ensure that problems in the performance history of the agent will be viewed more charitably than otherwise. For example, someone who felt that the UK government advisors act benevolently and are not swayed by interest groups and lobbies would have been more likely to have expressed trust in the government (and acted on the basis of its advice) after the incorrectness of its views on the safety of British beef became evident. However, within the TCC model, performance history of the agent should not influence trust in motives or factors that themselves influence trust in motives (i.e. judgments of similarity of values).

In their study, Harvey et al. (2007) varied the biases in consumer organizations' and government agencies' risk estimates. In their first experiment, these two sources of advice were, on average, both unbiased and equally (in)accurate. In their second experiment, they were again, on average, equally (in)accurate but the consumer organizations systematically overestimated risk and the government agencies systematically underestimated it to the same degree. Participants rated their values as less similar to those of both advisors in the second experiment. (Their stated trust in both advisors was also less in the second experiment.) Thus, activity in the trust-in-competence route (i.e. bias in performance history) influenced activity in the trust-

in-motives route (i.e. assessment of similarity of values). This should not occur according to the TCC model.

Social psychologists are familiar with self-enhancement effects (e.g. Krueger, 1998). People tend to overestimate the degree to which they possess traits that people regard as good and underestimate the degree to which they possess those that people regard as undesirable. Being biased in one's judgments is presumably considered to be undesirable. Hence people would be less likely to regard themselves as similar to people who make biased judgments than to those who do not. Thus self-enhancement provides a mechanism by which past history can influence activity in the trust-in-motives route.

If this proposal is correct, we should expect other aspects of the past history of an advisor's performance that are undesirable to have a similar effect to bias. For example, unreliability in the quality of advice is likely to be considered undesirable. Thus, all else being equal (including the average quality of advice), people should judge themselves to be more similar to those who produce advice of a consistent quality than to those who produce advice of variable quality.

Twyman, Harvey and Teki (2007) tested this hypothesis. In their experiments, all participants made risk assessments for a variety of hazards based on advice from two sources. Feedback about 'true' levels of risk enabled them to appreciate that one of their advisors (consumer organization or government agency) was better than the other. There were four conditions. In the two 'reliable-advisors' conditions, the good advisor and the bad advisor were of the same type throughout the experiment: thus, either the good advisor was a consumer organization and the poor advisor was a government agency or vice versa. In the two 'unreliable-advisors' conditions, the type of advisor producing a particular level of advice quality switched half way through the session. Thus, in one of these conditions, the consumer organization started as the good advisor and switched to being poor and the government agency started as being poor and switched to being good. In the other one, the opposite occurred: the consumer organization started out as the poor advisor and switched to being good whereas the government agency started as the good advisor and switched to being the poor one.

In line with the self-enhancement hypothesis, people assessed their values, interactions and goals to be more similar to reliable advisors than to unreliable ones. They also expressed more trust in reliable advisors – presumably because, as the TCC model suggests, they had judged these advisors to have values more similar to their own.

Implications for the TCC model

Two assumptions underlying the TCC model appear to require some modification. It would be relatively easy to eliminate the asymmetric influence assumption by incorporating a self-enhancement process by which aspects of an advisor's past performance affect how similar judges assess advisors' values to be to their own. The essential structure of the model would be preserved: indeed, this suggested modification is no more disruptive than the modification that Siegrist et al. (2003) themselves suggest on the basis of their own results from their study using electromagnetic field risks.

Dealing with the infringement of the final common pathway assumption is, perhaps, a little more problematic. The two routes of the TCC model are distinguished by the content of their processing (motives vs competence). We have suggested that the two-route models developed by cognitive psychologists that make a distinction between implicit and explicit processing may be relevant to trust in risk communication. However, those two-route models are distinguished by the manner of their processing (e.g. verbal/conscious vs nonverbal/nonconscious processing of the same content). It is not immediately clear to us what structure would result from merging these two rather different two-route models. However, it may be that processing of information related to trust in motives is primarily implicit and that processing of information related to trust in competence is largely explicit. This would simplify the resulting model but would require empirical support.

Implications for the study of trust

Trust is studied in different ways in different domains. Behavioural game theorists study trust-related behaviour (revealed trust) without asking participants to report their levels of trust in those with whom they are interacting. But in many negotiations expressions of trust (or lack of it) between those on the same side or on different sides may well influence trust-related behaviours. Behavioural game theory may be enriched and be given greater applicability by incorporation of statements of trust into some of the studies carried out within its framework.

Within risk communication, trust is studied by polls, surveys and questionnaires in which people are asked to express their trust using words, numbers or rating scales. This approach may inform us about the mental models or lay theories that people have of their trust-related behaviour (cf. Morgan, Fischhoff, Bostrum and Atman, 2002). However, if those mental models or lay theories lack validity, an approach based on studying stated trust alone may tell us little about their trust-related behaviour. If we want to know about trust-related behaviour, it is perhaps best to measure it directly.

References

- Baron, J., Hannan, M.T. and Burton, M.D. (2001). Labor pains: Changes in organizational models and employee turnover in young high-tech firms. *American Journal of Sociology*, 106, 960-1012.
- Berg, J.E., Dickhaut, J. and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10, 122-142.
- Berry, D.C. and Broadbent, D.E. (1984). On the relationship between task performance and associated verbalisable knowledge. *Quarterly Journal of Experimental Psychology*, 36, 209-221.
- Berry, D.C. and Broadbent, D.E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology*, 79, 251-272.
- Camerer, C.F. (2003). *Behavioral game theory: Experiments in strategic interaction*. New York: Russell Sage Foundation.
- Corritore, C.L., Kracher, B. and Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, 58, 737-758.
- De Bruin, E.N.M. and Van Lange, P.A.M. (1999). Impression formation and cooperative behaviour. *The European Journal of Social Psychology*, 29, 305-328.
- De Bruin, E.N.M. and Van Lange, P.A.M. (2000). What people look for in others: Influences of the perceiver and the perceived on information selection. *Personality and Social Psychology Bulletin*, 26, 206-219.
- Earle, T.C. and Cvetkovich, G. (1999). Social trust and culture in risk management. In G. Cvetkovich and R.E. Löfstedt (Eds), *Social Trust and the Management of Risk*. London: Earthscan (pp 9-21).
- Earle, T.C. and Siegrist, M. (2006). Morality information, performance information, and the distinction between trust and confidence. *Journal of Applied Social Psychology*, 36, 383-416.
- Fehr, E. and Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14, 159-181.
- Finucane, M.L., Alhakami, A., Slovic, P. and Johnson, S.M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Social Psychology*, 29, 305-328.
- Finucane, M.L. and Holup, J.L. (2006). Risk as value: Combining affect and analysis in risk judgments. *Journal of Risk Research*, 9, 141-164.
- Frewer, L.J., Howard, L., Hedderley, D. and Shepherd, R. (1996). What determines trust in information about food-related risks? Underlying psychological constructs. *Risk Analysis*, 16, 473-486.

- Fukuyama, F. (1995). *Trust: The Social Virtues and the Creation of Prosperity*. London: Penguin Books.
- Grabner-Kräutner, S. and Kaluscha, E.A. (2003). Empirical research in on-line trust: A review and critical assessment. *International Journal of Human-Computer Studies*, 58, 783-812.
- Harvey, N., Twyman, M. and Harries, C. (2007). Forms of trust: Interrelations, dissociations and compatibility effects. Manuscript under review.
- Kramer, R.M. and Tyler, T.R. (1996). *Trust in Organizations: Frontiers of Theory and Research*. Thousand Oaks, CA: Sage.
- Krueger, J. (1998). Enhancement bias in descriptions of self and others. *Personality and Social Psychology Bulletin*, 34, 505-516.
- Markova, I. (2004). *Trust and Democratic Transition in Post-Communist Europe*. Oxford: Oxford University Press.
- Mayer, R.C., Davis, J.H. and Schoorman, F.P. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20, 709-734.
- McAllister, D.J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 38, 24-59.
- Mettay, D. (1999). Institutional risk and confidence: A journey into a conceptual quagmire. In G. Cvetkovich and R. Löfstedt (Eds), *Social Trust and the Management of Risk*. London: Earthscan, pp 100-116.
- Morgan, M.G., Fischhoff, B., Bostrom, A. and Atman, C.J. (2002). *Risk Communication: A Mental Models Approach*. Cambridge: Cambridge University Press.
- O'Neill, O. (2002). *A Question of Trust*. Cambridge: Cambridge University Press.
- Pavlou, P.A. (2003). Consumer acceptance of electronic commerce – integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, 7, 69-103.
- Reber, R. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 5, 855-863.
- Renn, O. and Levine, D. (1991). Credibility and trust in risk communication. In R.E. Kasperson and P.J.M. Stallen (Eds), *Communicating Risks to the Public*. Dordrecht: Kluwer, pp 175-218.
- 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, 393-404.

- Samuelson, P.A. (1948). Consumption theory in terms of revealed preference. *Economia*, 15, 243-253.
- Siegrist, M., Earle, T. and Gutscher, H. (2003). Test of a trust and confidence model in the applied context of electromagnetic field (EMF) risks. *Risk Analysis*, 23, 705-716.
- Siegrist, M. Gutscher, H. and Earle, T.C. (2005). Perception of risk: the influence of general trust, and general confidence. *Journal of Risk Research*, 8, 145-156.
- Slovic, P., Finucane, M.L., Peters, E. and MacGregor, D.G. (2002). The affect heuristic. In T. Gilovich, D. Griffin and D. Kahneman (Eds). *Intuitive Judgment: Heuristics and Biases*. Cambridge: Cambridge University Press (pp. 397-420).
- Twyman, M., Harvey, N. and Teki, S. (2007). Effects of unreliability of advice on trust in sources of risk communication. Manuscript in preparation.
- Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175.

Address for correspondence:

Nigel Harvey

Department of Psychology

University College London

Gower Street

London WC1E 7BT

UK

Tel: +44 (0)20 7679 5387

Fax: +44 (0)20 7436 4276

Email: n.harvey@ucl.ac.uk

Running head: Two-route models of trust