Assessing Technology In The Absence Of Proof: Trust Based On The Interplay Of Others' Opinions And The Interaction Process

Peter W. de Vries

Dept. Psychology of Conflict, Risk, and Safety

University of Twente, Enschede, The Netherlands

Stéphanie M. van den Berg

Dept. Research Methodology, Measurement, and Data Analysis

University of Twente, Enschede, The Netherlands

Cees Midden

Dept. Human Technology Interaction

Eindhoven University of Technology, The Netherlands

Running head: Process feedback and system trust

Document type: Extended Multi-Phase Study (4 studies)

Word count: 12 969 (max. 13 500 for a 4-study article).

Structured abstract

Objective: The present research addresses the question how trust in systems is formed when unequivocal information about system accuracy and reliability is absent, and focuses on the interaction of indirect information (others' evaluations) and direct (experiential) information stemming from the interaction process.

Background: Trust in decision-supporting technology, such as route planners, is important for satisfactory user interactions. Little is known, however, about trust formation in the absence of outcome feedback, i.e. when users have not yet had opportunity to verify actual outcomes.

Method: Three experiments manipulated others' evaluations ("endorsement cues") and various forms of experience-based information ("process feedback") in interactions with a route planner, and measured resulting trust using rating scales and credits staked on the outcome. Subsequently, an overall analysis was conducted.

Results: Study 1 showed that effectiveness of endorsement cues on trust is moderated by mere process feedback. In Study 2, consistent (i.e., non-random) process feedback overruled the effect of endorsement cues on trust, whereas inconsistent process feedback did not. Study 3 showed that while the effects of consistent and inconsistent process feedback largely remained regardless of face validity, high face validity in process feedback caused higher trust than those with low face validity. An overall analysis confirmed these findings.

Conclusion: Experiential information impacts trust even if outcome feedback is not available, and, moreover, overrules indirect trust cues – depending on the nature of the former.

Application: Designing systems so that they allow novice users to make inferences about their inner workings may foster initial trust.

Keywords: system trust; process feedback; outcome feedback; consistency; face validity

Précis: System outcomes have been found to impact user trust. Less is known about trust formation when users cannot verify actual outcomes. Experiential information derived from the interaction

process impacts trust even if outcome feedback is absent, and overrules indirect trust cues – depending on whether the former appears consistent or random.

Assessing Technology In The Absence Of Proof: Trust Based On The Interplay Of Others' Opinions And The Interaction Process

Trust is generally acknowledged to play an important role in our interactions with technology, such as process automation, online applications, or consumer electronics. As with interpersonal trust, meaningful interaction requires sufficient levels of trust to enable reductions of uncertainty regarding the functioning of this particular system and its capabilities. Hence, the concept of system trust is crucial in understanding how people interact with systems, an idea that has firmly taken root in research in this field (for instance, see Halpin, Johnson, & Thornberry, 1973; Lee & Moray, 1992; Lee & See, 2004; Merritt, 2011; Muir, 1988; Sheridan & Hennessy, 1984; Verberne, Ham, & Midden, 2012; Zuboff, 1988).

Arguably, the antecedents of system trust depend, at least to some extent, on the degree of experience of the user. Someone who is experienced in using an online route planner, for instance, may base a trust judgement on his or her experiential information in terms of interaction outcomes, i.e., how often the system has provided advice that turned out to be correct. To the inexperienced user, the opinions and recommendations of others about a system are probably the easiest source of trust-relevant information, and, as such, they are influential in the user's decision to start using it (De Vries & Midden, 2008). As will be argued in the following sections, however, users may also gain direct experience even though actual outcome feedback is not available to them, for instance by simply test-running the application.

When it comes to direct experience (or direct information), the crucial distinction made in this paper is between outcome feedback and process feedback, or, in short, between feedback obtained from trying and testing a system and from test-running it. The availability of outcome feedback allows users to either verify a system's solutions or advice in terms of good or bad, or to decide to what extent they are satisfied with the provided advice. They may purchase an item online, and assess whether delivery was in conformance with what was promised beforehand. Similarly, a user may follow a route planner's driving directions and arrive at a particular final destination, and subsequently assess whether the suggested route's duration was indeed one hour and 35 minutes and whether traffic jams were successfully avoided. Process feedback, on the other hand, is used here to denote any kind of direct interaction in the absence of outcome feedback. Thus, people may try an online bookseller by entering a query for a particular book, adding the book to the shopping basket, acquiring information about shipping and handling costs, but stop the interaction before the deal is actually closed and outcome feedback may become available. Similarly, people seeking routing advice may try out a route planner by entering a few destinations and see what the system's suggestions will be without actually driving them. Thus, they actually engage in direct interaction with the application, even though outcome feedback is not yet available to them; after all, this would only be available after actually driving the suggested routes. Information obtained from process feedback does not necessarily have anything to do with actual algorithms and functions employed by the system (such as cost functions used by route planners to calculate routes) but is the result of the users' information processing based on the cues provided to them via a system's interface displays.

Recently there has been a marked increase in attention for other, more subtle trust cues in human-system interaction than outcome feedback, such as goal similarity (Verberne, Ham, & Midden, 2012) and cues conveying transparency and system rationale (e.g., De Visser *et al*. 2014; Helldin *et al*., 2013; Ososky *et al*., 2014; Thill, Hemeren, & Nilsson, 2014). Nevertheless, the effects on trust of direct experiences in the absence of outcome feedback have, to our knowledge, not received any attention in human factors research. The question central to this paper, therefore, is whether and how such direct experiences influence trust when feedback on the outcomes is absent, and how these interact with indirect information such as concurrently available recommendations of others.

Antecedents Of System Trust

System trust is defined here as a user's expectation about the system, that it will perform a certain task that is beneficial for the user, in a situation in which a lack of sufficient evidence causes the actual outcome of that task to be uncertain. It effectively limits the vast number of possible future interaction outcomes to only a relatively small number of expectations, thus reducing perceptions of both uncertainty and risk of the actor (Luhmann, 1979; cf. Giddens, 1990). Luhmann (1979) furthermore argued that trust should be seen as part of a continuous feedback loop that indicates whether or not trust is justified. More specifically, there is an object at which trust is directed, the referee or trustee, and this object provides feedback in terms of behaviour on the basis of which trust might be built up or broken down (cf. Lee & See, 2004). So, a system's behaviour may be watched by the user to see whether trust placed in it was justified. If the system performs according to the user's positive expectations trust may be maintained or increased; not living up to expectations will result in a breakdown of trust, possibly to the extent that trust is replaced by distrust. Luhmann's feedback loop emphasises the role of positive and negative interaction outcomes, i.e. direct information. These, however, are not available to novice users, who may have to rely on indirect information instead.

The effects of indirect information such as recommendations on trust have been studied in such diverse fields as consumer behaviour (Formisano, Olshavsky, and Tapp, 1982), reputation management (Standifird, 2001, Jensen, Davis, and Farnham, 2002), and web site credibility (Fogg and Tseng, 1999, Fogg *et al.,* 2001, Briggs, Burford, De Angeli, and Lynch, 2002), and it has been found to be of particular importance to trust in initial relationships (e.g., see McKnight, 2002; McKnight, 1998). System trust research, however, has largely neglected the role of indirect information (for an exception, see De Vries & Midden, 2008), and instead focusses on the build-up of trust as a function of personal experience over prolonged experimental trials. Typically, the focal system produces varying numbers of output errors, such as under- or overheating of juice or milk in a pasteurisation plant (Lee & Moray, 1994; Muir, 1989) or the incorrect classification of characters as either letters or digits (Riley, 1996), which are subsequently shown to influence trust and reliance on automation (also see De Vries, Midden, & Bouwhuis, 2003).

Such unequivocal output errors, however, may not be the only trust-relevant information obtainable from direct experience. Woods, Roth, and Bennett (1987), for instance, found that when technicians do not trust a decision aid, they either reject its solution to a problem or try to manipulate the output toward their own preconceived solutions. In their study, they found evidence that technicians, working with a system designed to diagnose faults in an electromagnetic device and suggest repairs, sometimes simply judged themselves whether the system′s pending advice was likely to solve the problem, rather than implementing the suggested change and subsequently checking whether it provided the desired results. In other words, these technicians apparently did not wait until unequivocal right/wrong feedback became available to them to form a trust judgement, but rather followed their own judgements on the plausibility of the system′s ″line of reasoning″ as it was fed back to them. Apparently, people sometimes judge the quality of system advice on the process that led to that advice.

Similarly, Lee and Moray (1992) argued that besides automation reliability, also ″process″ should be considered as a trust component of direct experiences. Process denotes an understanding of the system′s underlying functions or characteristics, such as the rules or algorithms that determine how the system behaves. As such, it bears resemblance to mental models, referring to representations that capture the workings or structure of a device (Sebrechts, Marsh, & Furstenburg, 1987). As such, they represent knowledge of how a system works, what components it consists of, how these are related, what the internal processes are, and how they affect components (Carroll & Olson, 1988). Mental models allow users to explain why a particular action produces specific results; however, they may be incomplete or internally inconsistent (Allen, 1997).

Such understanding of a system's inner workings may be facilitated by the degree of consistency of process feedback on which it is based. Analogous to interpersonal trust models, which hold that individuals are inferred to be dependable after they have consistently displayed instances of reliable behaviour (Rempel, Holmes, & Zanna, 1985), so too does making inferences about internal processes probably depend on consistency of system behaviour. Users may conclude there is a reason for the system′s process feedback to show a particular recurring pattern. For example, a user may request a route planner's advice on a number of different routes and subsequently notice that it persists in favouring routes that use a ring road over those that take a shortcut through the city centre. The user might then start conjecturing what causes this evident preference, and may, for instance, infer that the system may discard shortcuts through the centre because it is prone to dense traffic. Regardless of whether it actually matches the system′s actual decision rules, this insight in the system′s inner workings, comparable to, for instance, Zuboff's (1988) ″understanding″, Lee and Moray′s (1992) ″process″, and Rempel *et al.*′s (1985) ″dependability″, may reduce the user′s uncertainty, and, thus, lead to a greater willingness to rely on the system′s advice. Indeed, research by Dzindolet, Peterson, Pomranky, Pierce, and Beck (2003) has shown that participants working with a ″contrast detector″ to find camouflaged soldiers in terrain slides, trusted the system more, and were more likely to rely on its advice when they knew why the decision aid might sometimes fail, compared to those who were ignorant of such causes.

Although Dzindolet *et al.*′s (2003) studies provide additional, empirical support for the idea that a sense of understanding is beneficial for trust, their participants did not obtain this information from their own direct experiences with the device, as both Lee and Moray′s (1992) concept of ″process″ and mental model theory entails, but rather obtained it from the experimenter. As such, the assumption that users form such beliefs by observing system behaviour remains untested.

Combined effects of indirect and direct information

Normally, users probably have multiple concurrent types of information available to help them form a trust judgement about a particular system; besides their own experiences, based on process and outcome feedback, they may also resort to the opinions of others. Like accumulated prior experience with a system, such indirect information may influence users' perceptions of the system, and, hence, trust and automation use (cf. Merritt & Ilgen, 2008). Potentially important in this regard is the impact of both sources of information. Direct experiences have been argued to be more informative than indirect ones, and have been shown to lead to more robust attitudes (e.g., see Regan & Fazio, 1977). For the same reason, they have been argued to have a stronger influence on trust formation than indirect information (Arion, Numan, Pitariu, & Jorna, 1994). Congruously, Yuliver-Gavish (2011) showed that experiential information about a decision support system's performance had a stronger impact on users' reliance on the system than did descriptive information.

Arguably, whether or not direct experiences are superior to indirect experiences depends on the actual amount of information derived from these experiences. When a system′s process feedback is consistent, in that it displays stable preferences or patterns, this will allow users to generate a line of reasoning to explain the regularities. This type of feedback could therefore be considered as highly informative, and, as such, may be capable of overriding the influence of the less informative recommendations. Contrarily, inconsistent feedback may contain far less information that will be instrumental in the formation of such beliefs. As such, the information it conveys may not be substantial enough to override the effect of competing recommendations.

The current research

This section describes the results of three consecutive experiments and an overall analysis. All three experiments revolved around participants' interaction with a number of supposedly different route planners. The procedures for each off these experiments were largely identical; only the visual feedback about planned routes varied.

Outline of the studies

Study 1, a pilot study, was conducted to establish the influence of mere process feedback; specifically, we tested whether there would be a difference in the effect of with endorsement cues on system trust depending on presence or absence of process feedback, i.e. whether or not the generated routes would be visualised. Study 2 was designed to test the interaction of endorsement cues with a specific characteristics of process feedback, viz. its consistency; the set of routes displayed in Study 1 were adapted and supplemented to create a more homogenous set on the one hand and a set with a more jumbled appearance on the other. Specifically, in one condition, a stable preference for arterial roads or highways was displayed, whereas in the other routes were selected randomly from a subset of different routes. Study 3 aimed to partly replicate the findings of study 2 and simultaneously to extend it by disentangling the effect of consistency from that of face validity. In other words, this study tested the effect of consistency when the routes generated were high in face validity (as they were in Study 2) compared to when they were not, i.e. when they were unconvincing route options. Finally, Study 4 was conducted to further bolster the claim that user-system interaction provides trust-relevant information despite the absence of verifiable outcome feedback, an overall analysis was conducted combining the various manipulations in the three experiments, allowing us to assess the validity of the focal point with far greater statistical power.

Overall methodology

In all three experiments participants were seated behind a PC, where they were informed that they would participate in research concerning the way people deal with complex systems. Specifically, they would have to interact with four different route planners capable of determining an optimal route by estimating the effects of a vast number of factors, ranging from simple ones, like obstructions and one-way roads, to more complex ones, such as (rush-hour) traffic patterns. Furthermore, they were told that the computer had a database at its disposal, containing route information based on the reported long-time city traffic experiences of ambulance personnel and policemen from that city. These experiences supposedly constituted a reliable set of optimal routes, against which in principle both manually and automatically planned routes could be compared and subsequently scored; however, in these experiments only automatic route planning was enabled. As such, only the route planning capability of the machine was validated; the result of this validation, however, was fed back to participants only after completion of the entire experiment.

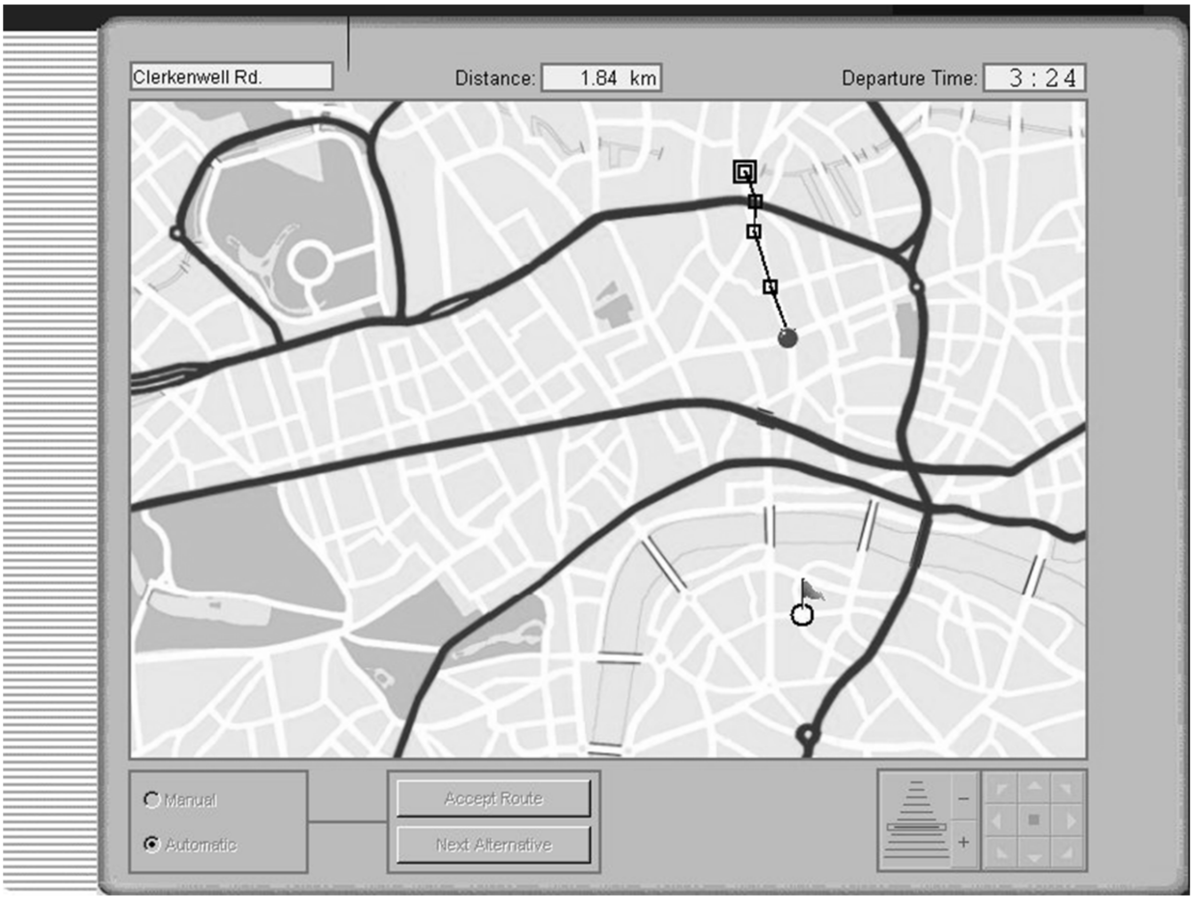

Figure 1. Route planner interface

During the experiments, a map was shown on the screen (see Figure 1); participants were not informed that it was based on the map of London. Using this map, participants were requested to perform a professional route dispatcher’s task by sending quickest possible routes to waiting cars, the current location and destination of which were indicated on the screen. The route-planning phase consisted of 5 trials with each of the four route planners; by clicking the “Automatic”-button the route-generating process was started. The automatically generated routes appeared on the screen in an incremental fashion, i.e. by drawing lines from each crossing to the next; the exact nature of the displayed routes varied between experiments and Process Feedback conditions. Finally, after the route had been generated the “Accept Route” button would become active; by clicking it the “dispatcher” supposedly sent the routing advice.

In all three experiments participants received information about the endorsement of the system by participants in a recent pilot test, and for each route planner this was either manipulated to be high or low (Endorsement Cue). Specifically, before actual interacting with each of the four route planners, high Endorsement Cue participants learned that a majority were extremely satisfied. In the low Endorsement Cue condition, participants were told that this was a minority. As all participants encountered both the low and high Endorsement Cues twice, two slightly different percentage figures were randomly used to convey high endorsement (“more than 83%” or “app. 88%”), and two for low endorsement (“less than 17%” or “app. 12%”).

We assumed participants would be more committed to the task if a certain risk were to be associated with their choices. Thus, we designed the experiment so that they were allotted ten credits per route-planning trial, which, either entirely or partially, could be put at stake. Directly after a route’s starting point and finish were indicated on the map, a dialogue box would appear on the screen, asking participants to enter any number of the allotted ten credits as stakes. The actual automatic route generation commenced immediately after they had entered this number. When an automatically generated route, after supposed comparison with the database with reported routes, was judged slower, participants would lose the credits they had staked on this particular route; a quicker route resulted in a doubling of the staked credits. Participants’ total number of credits would be revealed after interaction with all four route planners, and they were told that the money they would receive would depend on this total. However, as the program gave only bogus feedback, all participants were rewarded equally for their participation (€ 3. -, approximately US$ 3.50). Besides committing participants to their task, the number of credits that participants staked on the outcome of the automatic route-planning mode was considered a reflection of their trust in the system, with few staked credits indicating low trust, and many credits implying high trust (analogous to Berg, Dickhaut, & McCabe, 1995).

Both before and after interaction with each route planner, participants were required to rate the extent to which they trusted the system (7-point scales, ranging from “very little” (1.) to “very much” (7.)). Thus, we obtained self-reports of system trust, in addition to the measure of trust derived from the staking of credits.

In none of the studies outcome feedback, i.e., clear feedback in terms of a particular route being either successful or not, was made available to participants during interaction with the route planners.

Study 1 (pilot): Cue Effectiveness And Mere Process Feedback

Method.

Twenty-four undergraduate students (10 F, 14 M, M_{age} = 20.96, SD = 1.76, range = 18 - 24 y) participated in this study. The experiment had a 2 (Endorsement Cue: low versus high) * 2 (Process Feedback: present versus absent) within-participants full-factorial design.

In this study, participants were told that all route planners would generate routes but that some of them would and others would not actually visually present them (i.e., the Process Feedback present and Process Feedback absent conditions, respectively). Nevertheless, they were requested to stake credits and to accept each of these routes when the system indicated completion, i.e., when the “accept route” button would become active. The routes generated in the Process Feedback present condition were obtained in earlier experiments (reported in De Vries & Midden, 2008; De Vries, Midden & Bouwhuis, 2003), where participants could also manually plan routes. These manually planned routes were logged in a data file, from which the most commonly planned routes were selected to be used in this experiment. Thus, each presented route was deemed realistic by previous participants.

Results.

No effects were found for the order in which participants received the manipulations. Therefore, this variable will not be included in the subsequent analyses.

Before- and after-interaction trust measures.

A repeated-measures ANOVA was run with the trust ratings as dependent variable, and Endorsement Cue, Process Feedback, and Time of measurement (i.e., before versus after interaction) as independent variables. Means and standard deviations are displayed in Table 1.

Table 1
Average ratings of system trust, taken before and after interaction on 7-point scales, and standard deviations as a function of Endorsement Cue and Process Feedback; higher scores indicate higher levels of trust

	Trust Measure							
	Before				After			
	Process Feedback				Process Feedback			
	Present		Absent		Present		Absent	
Endorsement Cue	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	3.25	1.59	3.04	1.63	3.29	1.60	3.46	1.72
High	5.21	0.93	5.38	0.82	4.88	1.16	4.29	1.49

Both Endorsement Cue and Time of measurement produced (marginally) significant main effects, (F (1, 23) = 38.4; p < .01, and F (1, 23) = 3.6; p < .08, respectively). Process Feedback and the interaction between Endorsement Cue and Process Feedback did not yield significant effects, Fs < 1.

Endorsement Cue and Time of measurement, however, appeared to interact, F (1, 23) = 11.1; p < .01; the effect of the former was largest in the before-interaction measurements.

More interestingly, a significant three-way interaction between Endorsement Cue, Process Feedback and Time of measurement was found, F (1, 23) = 6.0; p < .03 (see Figure 2).

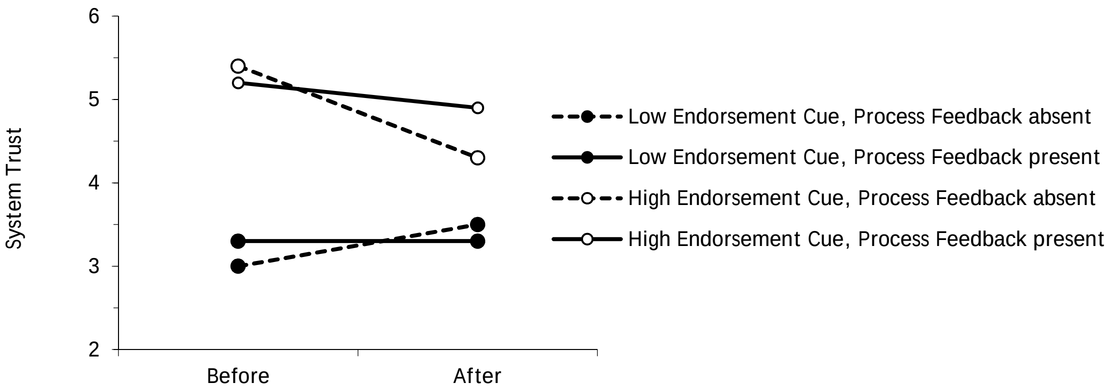

Figure 2. Average ratings of system trust, taken before and after interaction on 7-point scales, as a function of Endorsement Cue and Process Feedback; higher scores indicate higher levels of trust

Follow-up analyses showed that when Process Feedback was absent, Endorsement Cue and Time of measurement interacted significantly, F (1, 23) = 18.8; p < .01, indicating that the effect of

Endorsement Cue after interaction was less pronounced than before. When Process Feedback had been present, however, this interaction was non-significant, $F < 1$.

Staked credits.

The average number of credits staked was subjected to a repeated-measures ANOVA with Endorsement Cue and Process Feedback as independent variables. This analysis revealed a significant effect of Endorsement Cue, $F(1, 23) = 7.2$, $p < .02$ (sphericity assumed), indicating that participants had entered fewer credits in trials preceded by a low endorsement cue than in trials preceded a high endorsement cue. However, no significant effects of Process Feedback, or of an interaction were found, $F(1, 23) = 2.3$, *ns*., and $F(1, 23) < 1$. Results are shown in Table 2.

Table 2

Average number of staked credits and standard deviations as a function of Endorsement Cue and Process Feedback

		Process Feedback			
		Present		Absent	
		M	*SD*	*M*	*SD*
Endorsement Cue	Low	3.93	2.30	4.78	2.64
	High	5.38	2.41	5.68	2.45

The correlation between the number of credits staked and ratings of system trust was marginally significant, $r = .37$, $p < .08$.

Discussion.

The mere availability of process feedback proved to affect trust. The results showed that when no process feedback was given, the after-interaction trust measurements were less influenced by endorsement cues than before-interaction measurements. When process feedback was present, no such interaction was found. Whereas the former might be explained by the wearing-out of cue effectiveness over time, the latter could have been caused by the apparent randomness of the displayed routes. Somewhat jumbled visual information may have been difficult to interpret, and, thus, participants may have had to resort to cue content to support interpretation. This explanation would imply that less jumbled, i.e. more consistent process feedback would not invoke the need for cues for interpretation, as it may provide information, thus overruling rather than sustaining endorsement cue effects. This will be tested in Study 2.

Study 2: Cue Effectiveness And Process Feedback Consistency

This experiment was conducted to study the effects of endorsement information in combination with consistent versus inconsistent route generation.

Presumably, when a system′s feedback is consistent, it may enable users to generate beliefs about the system's workings that explain the regularities. As such, consistent process feedback could be considered to convey information. Contrarily, inconsistent process feedback may not convey such information. Consequently, consistency in the routes displayed on-screen while interacting with the route planner was expected to increase trust, whereas the absence of consistency, i.e., randomness, would have no such effect. In fact, as inconsistent process feedback may be interpreted as system inadequacy, it could be expected that an additional decrease in trust ratings would be found.

In the absence of process feedback, endorsement cues were expected to be used to form trust, as would become evident from the before-interaction trust measures. With the availability of process feedback, however, the information in the endorsement cue would have to compete with the information provided by process feedback. The process feedback characteristics would, therefore, determine what would happen to cue effectiveness. Specifically, the information conveyed by consistent process feedback was expected to override the influence of the competing, less informative endorsement cue on the after-interaction measures. The little information obtained from inconsistent process feedback, however, may not be substantial enough to override the effect of competing endorsement information. Consequently, when an inconsistent process determines the displayed route, the effect of an endorsement manipulation could be expected to be sustained over time, rather than overruled.

Method.

Thirty-two students participated in this study (6 F, 26 M, M_{age} = 22.06, *SD* = 1.81, range = 18 - 26 y). The experiment had a 2 (Endorsement Cue: low versus high) * 2 (Process Feedback: consistent versus inconsistent) within-participants full-factorial design.

In this study, the routes in both Process Feedback conditions were based on those used in study 1 and those in the log file with manually planned routes in earlier experiments (see De Vries & Midden, 2008; De Vries, Midden & Bouwhuis, 2003). This was done to keep face validity, i.e. the degree to which the routes were convincing as fastest routes or preferable in the eyes of participants, equal between the two conditions. For the consistent Process Feedback condition routes were selected that predominantly favoured arterial roads. Subsequently, sets of five different route alternatives were created for each combination of start and finish point; in the inconsistent Process Feedback condition the automatically generated route was randomly drawn from this set. As a result, routes in the Consistent Process Feedback condition took "red" roads, i.e. arterial roads or highways, in 80% of the cases, and deviated from the red routes in only 20 %. In the Inconsistent Process Feedback condition, the randomly selected roads either followed a red road in 20 % of the cases, whereas in the remaining 80 % a more-or-less straight line between start and finish or any other reasonably probable route was followed.

The manipulation checks required participants rate the extent to which (a) they could predict the generated routes, (b) they thought the generated routes displayed a certain pattern, (c) they thought that the generated routes were based on fixed rules, and (d) the generated routes matched the way they themselves would have planned them.

Results.

No effects were found for the order in which participants received the manipulations. This variable will, therefore, not be included in the subsequent analyses.

Manipulation checks.

Repeated-measures ANOVAs with Endorsement Cue and Process Feedback as independent variables showed that in the consistent Process Feedback condition (as opposed to the inconsistent condition) participants rated a higher ability to predict route generation, F (1, 31) = 44.3; $p < .01$, a greater extent to which they had discerned a certain pattern, F (1, 31) = 22.8; $p < .01$, a stronger belief that fixed rules were the basis for the generated routes, F (1, 31) = 15.3; $p < .01$, and a greater similarity of automatically generated routes with the way they themselves would have planned them, F (1, 31) = 8.3; $p < .01$. No effects of Consensus, nor of an interaction of Consensus and Process Feedback were found on any of these checks, all Fs ≤ 1.3; *ns.* The Process Feedback manipulation therefore proved successful.

Before- and after-interaction trust measures.

A repeated-measures ANOVA was performed, with Endorsement Cue, Process Feedback and Time of measurement (before- versus after-interaction) as independent variables. (See Table 3 for means and standard deviations).

Table 3

Average ratings of system trust, taken before and after interaction on 7-point scales, and standard deviations as a function of Endorsement Cue and Process Feedback; higher scores indicate higher levels of trust

	Trust Measure							
	Before				After			
	Process Feedback				Process Feedback			
	Consistent		Inconsistent		Consistent		Inconsistent	
Endorsement Cue	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	3.34	1.07	3.47	1.05	4.25	1.02	3.19	1.65
High	5.31	0.82	5.09	1.00	4.53	1.41	4.13	1.41

Several significant main effects were found. Trust was significantly higher after a high Endorsement Cue than after a low Endorsement Cue, F (1, 31) = 38.1; p < .01.; additionally, consistent Process Feedback resulted in higher trust than inconsistent Process Feedback, F (1, 31) = 6.7; p < .02. Time of measurement also yielded a significant overall effect on trust, F (1, 31) = 4.4; p < .05; overall, trust levels tended to decrease over time. No interaction between Endorsement Cue and Process Feedback was found, F (1, 31) = 0.4, *ns*.

The effect of Endorsement Cue was more pronounced on the before-interaction than on the after-interaction measure, as indicated by a significant interaction of Endorsement Cue and Time of measurement, F (1, 31) = 17.5; p < .01. Moreover, a significant three-way interaction between Process Feedback, Endorsement Cue and Time of measurement, F (1, 31) = 4.4; p < .04 was found. This interaction is visualised in Figure 3.

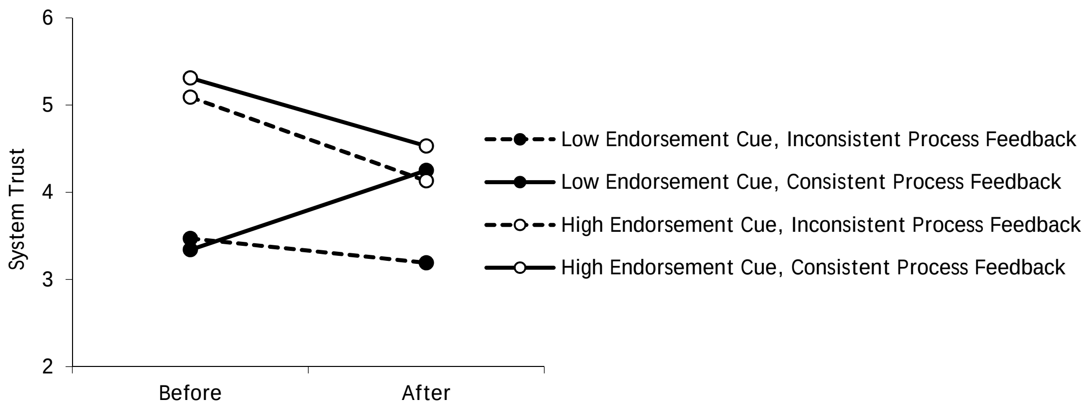

Figure 3. Average ratings of system trust, taken before and after interaction on 7-point scales, as a function of Endorsement Cue and Process Feedback; higher scores indicate higher levels of trust

Follow-up analyses were conducted to test the specific hypotheses pertaining to this three-way interaction. When Process Feedback was random, before- and after-interaction measures were both significantly affected by the Endorsement Cue manipulation, $F\ (1, 31) = 47.2$; $p < .01$, and $F\ (1, 31) = 5.72$; $p < .03$, respectively; as can be seen in Table 3, trust ratings were significantly higher following a high Endorsement Cue than they were after a low Endorsement Cue. In the Consistent Process Feedback condition, a highly significant interaction of Endorsement Cue and Time of measurement was found, $F\ (1, 31) = 25.5$; $p < .01$, indicating that the Endorsement Cue manipulations only had an effect on the before-interaction trust measurement ($F\ (1, 31) = 63.1$; $p < .01$), but not on the after-interaction measurement ($F\ (1, 31) = 0.9$; *ns.*).

These analyses therefore provided support for the hypothesis that inconsistent Process Feedback caused the Endorsement Cue effect to be sustained over time, whereas consistent Process Feedback overruled the effect of Endorsement Cue.

Staked credits.

The number of stakes entered showed a significant effect of Endorsement Cue, $F\ (1, 31) = 5.3$; $p < .03$. A high Endorsement Cue caused participants to stake more credits than a low Endorsement Cue. Process Feedback did not produce a significant effect, $F < 1$, *ns*. The interaction between Endorsement Cue and Process Feedback was not significant at the 0.05-level, $F\ (1, 31) = 3.1$, $p = .09$. See Table 4.

The ratings of system trust and the average number of staked credits correlated significantly, $r = .37$, $p < .04$.

Table 4
Average number of staked credits and standard deviations as a function of Endorsement Cue and Process Feedback

	Process Feedback			
	Consistent		Inconsistent	
Endorsement Cue	*M*	*SD*	*M*	*SD*
Low	5.00	2.17	5.18	2.25
High	5.98	1.67	5.48	1.97

Discussion.

The data showed that the differences in trust between high and low endorsement treatments hardly changed over time when process feedback was of a rather random nature, an indication that inconsistent process feedback provided little additional information that competed with endorsement information. In addition, participants may also have used the endorsement information to interpret the ambiguous randomised information presented on the screen. In the consistent process feedback treatments a different pattern emerged. Although endorsement information influenced participants' before interaction trust levels, this effect could not be shown for the after-interaction measure, which was in line with the hypotheses.

Both the trust measures and the credits staked were influenced by the endorsement information. Contrary to the trust measures, however, the credits did not show a reduction in this effect as a result of consistent process feedback. An explanation for this marked difference could lie in differences in "exposure duration" between endorsement cues and process feedback manipulations. The former was administered before participants started their interaction with each route planner, and, thus, could well have affected the credits staked in the all trials, including the first few. How consistent or inconsistent its process feedback was, on the other hand, could only be assessed after at least a few, and perhaps all five trials. Consequently, the effect of process feedback may simply not have been strong enough to manifest itself in the average over all five trials.

This experiment showed that the character of the process feedback plays a significant role in the formation of trust. One explanation for this finding, suggested previously, is that, contrary to randomness, consistency tempts users to think that there is a reason why the route planner results showed a particular recurrent pattern, rather than consider the pattern as an imperfection of the system. In other words, users may form beliefs about the system's functioning in order to explain its output. This, in turn, may increase trust and, subsequently, the willingness to rely on generated route solutions. The findings that the effect of endorsement information depended on the consistent versus inconsistent appearance of the suggested routes, and that participants were more convinced that there were fixed rules embedded in the route planners that gave consistent process feedback than those with inconsistent process feedback, provides additional support for this contention.

Study 3: Process Feedback Consistency And Face Validity

Arguably, consistency alone may not provide sufficient grounds for trust to form; users may also base their judgement on face validity. Indeed, one may think of a system yielding output that consists of consistent yet unlikely, or disagreeable advice. Being based on manually planned routes in earlier studies the consistent and inconsistent process feedback likely consisted of rather agreeable routing advice; the question remains what the influence of consistency will be when the routes displayed are highly unlikely as correct solutions, i.e., when routes are low in face validity and users are not likely to agree with the advice given to them.

Lerch and Prietula (1989) investigated agreement with human and system advice, and confidence in the source of this advice. They treated participants' agreement with system advice as similar to predictability, and proposed an additive model of confidence and agreement. Agreement ratings, like predictability ratings, were primarily guided by the specific evidence provided in each problem solving trial; confidence levels were to a certain extent based upon prior confidence levels and on an agreement history. By considering agreement similar to consistency, Lerch and Prietula implied that both concepts have a similar direct relation to trust. Higher agreement with system advice corresponds to higher levels of trust, as would be the case for consistency.

Face validity, or agreement, in Lerch and Prietula's (1989) terminology, and consistency in process feedback come about differently, however. Face validity of system advice, i.e., the extent to which people regard the advice as realistic, convincing, or preferable, may be based on one single route-planning trial; contrarily, conclusions concerning the degree of consistency can only be drawn after viewing multiple different routes. In other words, to a novice user, an assessment of face validity may be made before process feedback is judged as consistent. Furthermore, consistency does not necessarily imply that users agree with it. For example, if a user wants advice on how to travel from, say, the Royal Albert Hall to Piccadilly Circus, and subsequently from Piccadilly Circus to Tower Bridge, a route planner that consistently incorporates the distant Hyde Park in its suggestions is not very likely to instil trust in the user, and is probably not considered to provide feedback high in face validity. Therefore, consistency and face validity can be considered as separate characteristics of process feedback, and they will be treated accordingly in this study.

In Study 3, face validity of process feedback was pitted against process feedback consistency and endorsement cues. Similar to Study 2, consistent feedback was expected to result in higher trust ratings than inconsistent feedback. Likewise, process feedback with high face validity, i.e., process feedback that participants believe is likely to result in fast routes, would cause trust ratings to be higher than process feedback with low face validity, i.e., that is unlikely to yield fast routes. Additionally, as consistent process feedback may contain trust-relevant information, it was expected to overrule the effect of endorsement information, causing the effect of endorsement on the after-interaction trust measures to disappear. Inconsistent process feedback, being low in informational content, would not be able to overrule endorsement information, as would be indicated by a sustained endorsement effect on after-interaction trust over time.

As a result of the overruled endorsement effect, trust levels in the consistent conditions would show a convergence over time, as was observed in Study 2; whether trust levels would converge on high or low after-interaction trust levels, was expected to depend on face validity. Specifically, consistent process feedback with high face validity was expected to converge at higher trust levels than consistent process feedback with low face validity. Inconsistent process feedback was expected to show a sustained effect of endorsement on the after-interaction measure, in addition to an effect of face validity: inconsistent process feedback with high face validity would result in higher after-interaction trust than inconsistent process feedback with low face validity.

Method.

Participants and design.

Forty-eight undergraduate students participated in this study (9 F, 39 M, M_{age} = 21.69, SD = 1.81, range = 18 - 29 y), which had three-factor mixed design (full-factorial). Endorsement Cue (low versus high) was varied between-participants, whereas Consistency (consistent versus random) and Face Validity (high versus low) were manipulated within-participants. The order in which the Face Validity conditions were encountered constituted an additional two-levels between-participants variable.

Procedure.

Process feedback that was Consistent and had High Face Validity consisted of routes that favoured arterial roads, and, as such, were similar to the routes used in the consistent process feedback condition of Study 2. Likewise, routes displayed in the Inconsistent and High Face Validity conditions were the same as those used as inconsistent routes in the previous experiment, and showed routes selected randomly from a small subset of alternatives that participants had preferred in earlier experiments. Contrarily, the routes in the Low Face Validity condition, both Consistent and Inconsistent, were entirely different to the routes used before. Low Face Validity entailed routes that displayed relatively large detours; these routes, therefore, were not very likely to be as fast as required. Process feedback that was Consistent and had Low Face Validity showed routes that made a relatively large detour that was always on the same location; thus, these routes were both unlikely to be fast, but at the same time displayed consistency. Contrarily, process feedback that was Inconsistent and had Low Face Validity consisted of routes that made relatively large and inconsistent detours, i.e., never on the same spot.

The order in which these manipulations took place was counterbalanced. The first two route planners yielded Consistent process feedback, whereas the third and fourth were random, and vice versa. Within the Consistent and Inconsistent conditions, High and Low Face Validity conditions were systematically varied.

The manipulation checks concerning Face Validity entailed asking participants to rate the extent to which (a) the generated routes matched the way they themselves would have planned them, and (b) they agreed with the displayed routes. Consistency manipulations were checked by having participants rate the extent to which they (a) could predict the generated routes, (b) thought the generated routes displayed a certain pattern, and (c) thought that the generated routes were based on fixed rules.

Results.

The order in which manipulations in process feedback were encountered, proved to influence some dependent variables. The variable Order was, therefore, included in all reported analyses as an extra independent variable; as such, the reported effects are corrected for order effects. As no specific hypotheses regarding order effects have been formulated, they will only be discussed briefly where relevant.

Manipulation checks.

All manipulation checks were subjected to an ANOVA, with Consistency and Face Validity as within-participants independent variables, and Endorsement Cue and Order as between-participants independent variables.

The two checks concerning the extent to which the generated routes matched the way participants would have planned them themselves (similarity ratings), and the extent of agreement with the displayed routes both showed highly significant effects of Face Validity, $F\ (1, 32) = 43.8$, $p < .01$, and $F\ (1, 32) = 36.9$, $p < .01$, respectively. Ratings with regard to the former check were higher in case of High Face Validity than in case of Low Face Validity ($M = 5.46$, $SD = 2.16$ versus $M = 4.31$, $SD = 2.23$ in the Consistent condition, and $M = 4.90$, $SD = 2.22$ versus $M = 3.00$, $SD = 2.34$ in the Inconsistent condition). A similar effect of Face Validity was found on the latter check ($M = 5.63$, $SD = 2.05$ versus $M = 4.67$, $SD = 2.06$ in the Consistent condition, and $M = 5.29$, $SD = 2.20$ versus $M = 3.31$, $SD = 2.24$ in the Inconsistent condition). Both checks, however, also showed an effect of Consistency, $F\ (1, 32) = 12.4$; $p < .01$, and $F\ (1, 32) = 22.2$; $p < .01$. As can be observed above, both ratings were highest in the Consistent condition. In addition, a significant interaction between both independent variables was found on the agreement rating, $F\ (1, 32) = 5.1$; $p = .03$. It appeared that larger differences between High Face Validity and Low Face Validity were found in the Consistent conditions.

Furthermore, analysis of the Consistency manipulation check showed that participants judged the process feedback as significantly more predictable in the Consistent condition, compared to the

Inconsistent condition (M = 6.21, SD = 2.21 versus M = 4.65, SD = 2.36 in the Consistent condition, and M = 4.88, SD = 2.19 versus M = 2.96, SD = 2.41 in the Inconsistent condition), F (1, 32) = 42.4, p < .01. Also, a highly significant effect of Face Validity became apparent on this check, F (1, 32) = 56.7, p < .01; predictability was rated higher when process feedback had been high in Face Validity, versus when Face Validity had been low.

Consistency appeared to have a similar effect on the check to what extent participants had discerned patterns in the process feedback (M = 6.56, SD = 2.31 versus M = 5.50, SD = 2.40 in the Consistent condition, and M = 5.88, SD = 1.97 versus M = 4.69, SD = 2.59 in the Inconsistent condition), F (1, 32) = 7.7, p < .01, as did Face Validity, F (1, 32) = 15.6, p < .01.

Ratings regarding the extent to which they believed fixed rules to underlie system output showed only a marginally significant effect of Consistency, with higher scores in the Consistent condition, compared to the Inconsistent condition (M = 6.79, SD = 1.88 versus M = 5.50, SD = 2.13 in the Consistent condition, and M = 6.06, SD = 1.73 versus M = 5.38, SD = 2.38 in the Inconsistent condition), F (1, 32) = 3.0, p = .09. The effect of Face Validity, with High Face Validity resulting in higher scores than Low Face Validity, was significant, F (1, 32) = 11.8, p < .01.

Before and after-interaction trust measures.

The before- and after-interaction trust measures were subjected to ANOVAs, with Consistency and Face Validity as within-participants independent variables, and Endorsement Cue and Order as between-participants independent variables. Table 5 and Table 6 display means and standard deviations of before- and after-interaction trust ratings, respectively.

Table 5
Average ratings of system trust, taken before interaction on 7-point scales, and standard deviations as a function of Endorsement Cue, Consistency, and Face Validity; higher scores indicate higher levels of trust

		Consistency			
		Consistent		Inconsistent	
Face Validity	Endorsement Cue	*M*	*SD*	*M*	*SD*
High	Low	3.88	1.23	3.67	1.46
	High	4.96	1.12	4.71	1.16
	Total	4.42	1.29	4.19	1.41
Low	Low	3.88	0.99	4.21	1.02
	High	4.83	0.96	4.88	0.99
	Total	4.35	1.08	4.54	1.05

Table 6
Average ratings of system trust, taken after interaction on 7-point scales, and standard deviations as a function of Endorsement Cue, Consistency, and Face Validity; higher scores indicate higher levels of trust

		Consistency			
		Consistent		Inconsistent	
Face Validity	Endorsement Cue	*M*	*SD*	*M*	*SD*
High	Low	4.29	1.08	4.54	1.32
	High	4.67	1.43	5.00	1.10
	Total	4.48	1.27	4.77	1.22
Low	Low	3.71	1.55	1.92	1.06
	High	3.50	1.69	2.38	1.47
	Total	3.60	1.61	2.15	1.29

The effect of the Endorsement Cue manipulation significantly affected before-interaction trust measures, but not the after-interaction measures, F (1, 32) = 15.0; $p < .01$, and F (1, 32) = 1.4; *ns.* Trust was rated higher when a high endorsement cue was given, compared to a low endorsement cue. Apparently, the manipulations that took place in the interaction stage, i.e., after the before-interaction trust measure, overruled the effect of the Endorsement Cue.

The after-interaction measures showed a significant main effect of Consistency, F (1, 32) = 22.2; $p < .01$, indicating that these ratings were higher in the Consistent than in the Inconsistent conditions. Manipulations of Face Validity also affected the after-interaction trust measures; these were higher in the High Face Validity conditions than in the Low Face Validity conditions, as indicated by a significant main effect of Face Validity, F (1, 32) = 110.7; $p < .01$.

These results supported the hypotheses. When no other information was available, the endorsement information was used to build trust, as indicated by the Endorsement Cue effect on the before-interaction trust measures. After the interaction, however, Endorsement Cue no longer showed an effect on trust, as it was overruled by the competing information conveyed by process feedback.

The interaction of Face Validity and Consistency was significant for the after-interaction trust measures, F (1, 32) = 29.3; $p < .01$. Table 6 and Figure 4 show that that the effect of Face Validity was far smaller when Process Feedback was also consistent, compared to when it was random. The interaction on the after-interaction measures, however, indicates that Consistency was more influential than Face Validity. When Process Feedback was consistent, the fact whether it also had High or Low Face Validity added only little in terms of trust. Face Validity gained in importance in the absence of consistency, however, arguably as it did not have to compete.

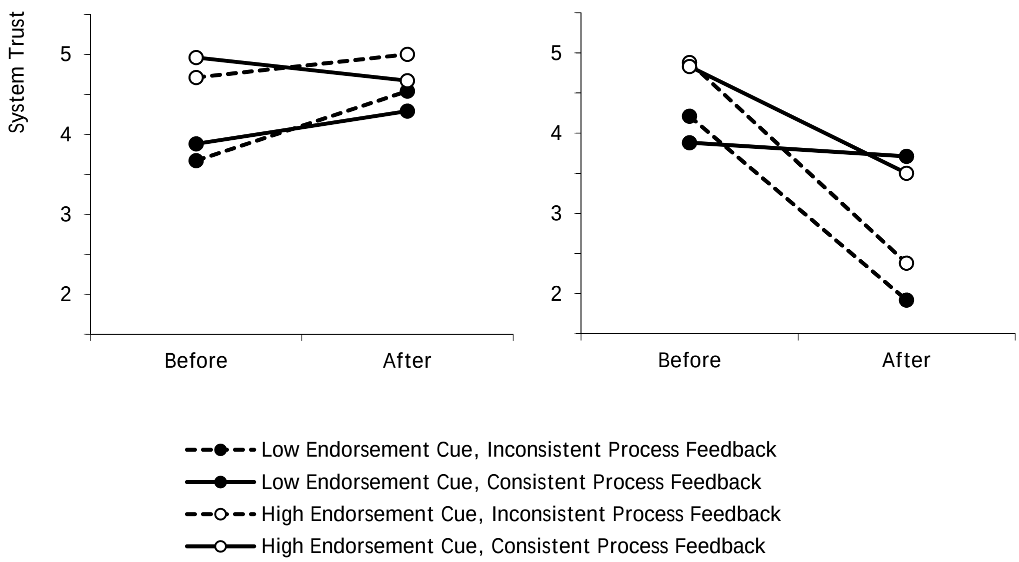

Figure 4. Average ratings of system trust, taken before and after interaction on 7-point scales, as a function of Endorsement Cue and Consistency; the left part shows averages for High Face Validity, the right part for Low Face Validity; higher scores indicate higher levels of trust

To test the specific hypotheses about the dependence of cue effectiveness on whether Process Feedback was consistent or random, separate analyses were run for Consistent and Inconsistent conditions. In the Consistent condition, a highly significant interaction between Time of measurement and Endorsement Cue was found, F (1, 32) = 11.8; $p < .01$; as expected, the effect of Endorsement Cue reached significance only for the before-interaction-, and not the after-interaction measures, F (1, 32) = 18.3; $p < .01$, and F (1, 32) = 0.1; *ns.*, respectively. A non-significant three-way interaction between Time of measurement, Endorsement Cue and Face Validity indicated that this effect could not be shown to differ between the High and Low Face Validity conditions, F (1, 32) = 0.5; *ns.* This supported the hypothesis that consistent Process Feedback would yield trust-relevant information that would overrule the competing, less informative Endorsement Cue.

In the Inconsistent condition, the interaction between Time of measurement and Endorsement Cue was not significant, F (1, 32) = 1.9; *ns.* Closer inspection revealed a significant Endorsement Cue-effect on the before-interaction-, and a marginally significant effect on the after-interaction measure, F (1, 32) = 9.4; p < .01, and F (1, 20) = 3.6; p = .07. A non-significant three-way interaction between Time of measurement, Endorsement Cue and Face Validity suggested this not to differ between High and Low Face Validity conditions, F (1, 32) = 0.6; *ns.* Although marginally significant, the after-interaction trust ratings showed an effect of the Endorsement Cue manipulation, which is in conformance with expectations: as inconsistent Process Feedback would convey only little competing trust-relevant information, the effect of Endorsement Cue information was expected to affect both before- and after-interaction measures.

The Order in which Process Feedback manipulations took place, appeared to interact with Consistency on the after-interaction trust measures, F (7, 32) = 5.4; p < .01. Subsequent analyses indicated that after-interaction trust ratings were somewhat higher when participants had encountered inconsistent Process Feedback first. In addition, the effect of Consistency manipulations on trust appeared to be strongest when inconsistent preceded consistent routes. Perhaps, when inconsistent Process Feedback was encountered first, the subsequent consistent routes may have been more easily recognisable as such, resulting in higher trust ratings following consistent Process Feedback, compared to when consistent routes were encountered first.

Staked credits.

No significant between-participants main effect of Endorsement Cue was found on the number of credits staked, F (1, 32) < .1, *ns*. Consistency only resulted in a marginally significant main effect, F (1, 32) = 3.7, p = .06. The number of credits staked was slightly higher in the consistent process feedback condition than in the inconsistent condition (see Table 7).

Contrarily, a highly significant main effect of Face Validity was found, F (1, 32) = 37.9, p < .01; high Face Validity caused participants to stake more credits than low Face Validity.

Table 7
Average number of staked credits and standard deviations as a function of Endorsement Cue, Consistency, and Face Validity

		Consistency			
		Consistent		Inconsistent	
Face Validity	Endorsement Cue	*M*	*SD*	*M*	*SD*
High	Low	5.29	1.99	5.25	1.80
	High	5.48	2.21	5.76	2.13
	Total	5.39	2.08	5.50	1.97
Low	Low	4.75	2.14	3.97	1.97
	High	4.43	2.50	3.70	2.49
	Total	4.59	2.31	3.83	2.22

Moreover, Face Validity and Consistency were found to interact significantly, F (1, 32) = 6.3, p = .02; as is illustrated by Figure 5, in the Consistent conditions, the manipulations of Face Validity turned out to have a smaller effect than in the Inconsistent condition.

The correlation between the number of staked credits and the system trust ratings was highly significant, r = .49, p < .01.

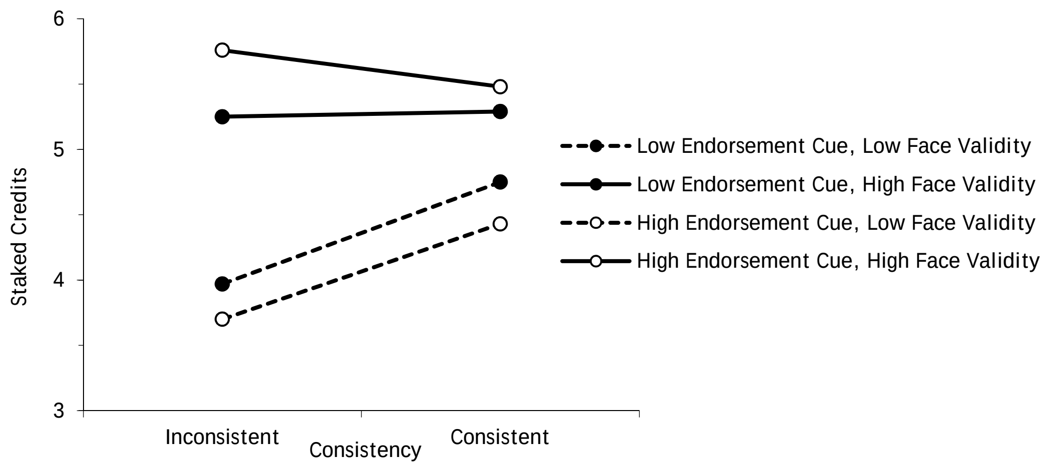

Figure 5. Average number of staked credits, as a function of Endorsement Cue, Consistency, and Face Validity

Additional analyses.

One could argue that the effect of process feedback is not so much the result of its consistency conveying information, but rather of its consistency simply being more preferable to users. To address this potential explanation, a hierarchical regression was conducted in which Consistency, Face

Validity, and their interaction term were inserted as predictors in the first model, and the agreement and similarity ratings that were part of the manipulation checks as additional predictors in the second model; this was done for both after-interaction trust and staked credits as dependent variables. As Consistency, in contrast to Face Validity, was expected to develop over trials, the staked credits of only the final (fifth) route-planning trial was inserted as dependent variable. As can be seen in Table 8, the addition of agreement and similarity to the second model did not at all change the magnitude and significance of the relationships of Consistency, Face Validity and their interaction with both dependent variables in the first model. These results, therefore, show that the effects of the independent variables and their interaction on both after-interaction trust ratings and the number of staked credits cannot be explained by similarity and agreement ratings.

Table 8
Results of a hierarchical regression

	After-interaction trust			Credit staked in 5th trial		
	β	*t*	*p*	*β*	*t*	*p*
1. Consistency	0.26	1.41	.16	0.42	1.92	.06
Face Validity	0.61	3.30	< .01	0.59	2.70	< .01
C x FV	1.13	4.47	< .01	0.94	3.12	< .01
2. Consistency	0.26	1.59	.11	0.45	2.20	.03
Face Validity	0.56	3.37	< .01	0.57	2.79	< .01
C x FV	0.95	4.16	< .01	0.79	2.82	< .01
Similar	0.09	0.73	.47	0.34	2.27	.02
Agree	0.31	2.58	.01	0.06	0.41	.68

Discussion.

The analyses reported here showed that, in line with previous experiments, endorsement information affected before-interaction trust levels and its effect on after-interaction trust actually depended on the nature of the generated routes. Apparently, depending on their nature, the routes displayed during the interaction stage provided participants with information that overruled the endorsement effect on the subsequent after-interaction trust ratings. As hypothesised, displayed routes that were likely to be fast routes (i.e., routes with high face validity), resulted in higher levels of trust than did routes that were unlikely to be fast (routes with low face validity). In conformance with the expectations, routes that were consistent were shown to cause higher trust ratings and higher numbers of staked credits than inconsistent routes. Interestingly, consistency and face validity also appeared to interact with one another: face validity proved to have a stronger influence when process feedback was also random, compared to when it was consistent.

With regard to the relation between cue effectiveness and consistency in process feedback, the analyses show that, in accordance with the specific hypotheses, consistent process feedback condition caused the endorsement manipulation to affect only the before-interaction, and not after-interaction trust levels. In other words, cue effectiveness was shown to be cancelled out over time when process feedback had been consistent. This effect did not differ between high and low face validity conditions. Contrarily, in the inconsistent process feedback condition, endorsement cues affected both before- and after-interaction trust, and this effect was visible in both face validity conditions.

This experiment showed that, besides consistency, face validity of the displayed routes also has an influence on trust. Process feedback with high face validity, or the displaying of routes that seemed likely to be fast, matched participants′ preconceptions about fast routes, and, thus, influenced trust. Likely fast routes resulted in higher trust levels than did unlikely fast routes (i.e., process feedback with low face validity). However, as noted in the above, the magnitude of the effect was determined by consistency. This could be interpreted as consistency having a higher ″priority″ than face validity; it seems as if participants rely more heavily on face validity when consistency is absent.

Lerch and Prietula (1989) reasoned that both agreement and confidence are rooted in predictability, or consistency, and, hence, are directly related. This explanation, however, fails to explain the interaction effects found on the after-interaction trust measurement and the number of staked credits. If predictability, or consistency, and face validity, or agreement, were linked as proposed by Lerch and Prietula, one would expect these to be additive. In other words, only main effects of these variables on both trust measures and the number of stakes would have been expected, but not an interaction. Significant interactions between consistency and face were found, however, indicating that these variables are not directly related as implied by Lerch and Prietula (1989). A consistent set of routes could indeed be judged as higher in face validity, but face validity does not necessitate consistency, as this experiment shows. In other words, consistency may directly influence face validity and trust, but face validity may also affect trust without consistency.

Relatedly, one could argue that the effects of process feedback could have more to do with user possible preferences for routes of a consistent nature than with consistency causing users to infer rules or information. Unfortunately, no direct measure was available to unequivocally support the proposed rule-inference mechanism. Nevertheless, the alternative preference explanation is not supported by the results presented here. Specifically, a hierarchical regression showed that the effects of the manipulation in this study were not affected by inclusion of measures tapping into participants' preferences, indicating that their effects are independent of these preferences.

Study 4: Overall Analysis

An overall analysis was conducted to provide further support for our main point that, despite the absence of verifiable outcome feedback, the visual process feedback a system provides trust-relevant information, and that consistency and face validity are instrumental and independent elements in this feedback. This analysis compared the effects of the various manipulations described in this paper across experiments 1, 2, and 3, thus allowing us to assess the validity of the focal point with far greater statistical power. To do so, the experimental conditions that were identical across the experiments were identified and combined.

Endorsement information was manipulated similarly across all three experiments, apart from the fact that in Study 3 manipulations took place between-participants, whereas in Studies 1 and 2 Endorsement was manipulated within-participants. All other variables were manipulated within-participants. Thus, for each participant there are four measurements taken before the interaction (i.e., one for each of the four route planners), showing only an effect of Endorsement manipulations, and four measurements taken afterwards, on which the process feedback manipulations had an additional effect.

The process feedback as manipulated in Study 1 was based on manually planned routes logged in earlier experiments (De Vries & Midden, 2008; De Vries, Midden & Bouwhuis, 2003) and provided the basis for the manipulations in Studies 2 and 3; routes in Study 1 and the log file were selected to create process feedback with a more inconsistent appearance in the one condition, and more consistent in the other. The process feedback supplied in Study 1 can therefore be considered to be between the consistent and inconsistent process feedback conditions of Study 2 in terms of consistency.

In Study 3, an additional characteristic of process feedback was added to the consistent versus inconsistent appearance of the routes used in Study 2, namely whether the displayed routes were likely to be fast routes (high face validity) or not (low face validity). In other words, routes were created that contrasted with the other process feedback manipulations to the degree that they were likely to yield successful routes; compared to these conditions, the other manipulations, i.e. the available condition in Study 1 and the consistent and inconsistent conditions in Study 2, can, therefore, be considered to have high face validity.

Table 9 shows how process feedback conditions of Studies 1, 2, and 3 are combined to form the conditions in the overall analysis. This analysis comprised levels in which process feedback Consistency can be unavailable (Absent Process Feedback), inconsistent (Inconsistent Process Feedback), available and in-between consistent and inconsistent (Available Process Feedback), and consistent (Consistent Process Feedback). In addition, process feedback can either have High or Low Face Validity.

As the process feedback manipulations of Study 2 were identical to the manipulations of Consistency in the High Face Validity condition of Study 3, these conditions can be combined. As such, the overall design consisted of only six different manipulations of Process Feedback (see Table 9). Taking the two levels of the Endorsement manipulation into account, combining all three experiments resulted in a design of 12 cells.

Table 9

Process feedback conditions in separate experiments compared to those in the overall analysis

	Separate Experiments	Overall Analysis
Study 1	Absent	Absent
	Present	Present - High Face Validity
Study 2	Consistent	Consistent - High Face Validity[a]
	Inconsistent	Inconsistent - High Face Validity[b]
Study 3	Consistent-High Face Validity	Consistent - High Face Validity[a]
	Inconsistent-High Face Validity	Inconsistent - High Face Validity[b]
	Consistent-Low Face Validity	Consistent - Low Face Validity
	Inconsistent-Low Face Validity	Inconsistent - Low Face Validity

Note. Indexes [a] and [b] indicate identical conditions that are combined in the overall analysis

Trust measurements taken before interaction depended only on the Endorsement manipulation, on individual differences in trust, and on within-participant variability. Thus, before-interaction trust (T) of an individual *i* after Endorsement manipulation *j* was modelled as a weighted sum of a fixed endorsement effect μ_j, a random variable A_i for individual differences in general trust, and a random variable E_{ij} for measurement error:

$$T_{ij,\text{ before}} = \mu_j + \alpha * A_i + \eta_j * E_{ij}, \qquad \text{where A and E} \sim N(0, 1).$$

Note that the model allowed for different error variances for the two types of endorsement. Subsequently, both an additive, as well as an interactive model was fitted to these data. The additive model was used to determine whether there are main effects of Endorsement Cue, Consistency and Face Validity manipulations, and what the magnitudes of the individual factors' effects are in terms of trust. Subsequently, comparing the additive and the interactive model with regard to how well each accounts for the observed means, yields information about interaction effects; if the additive model would provide a significantly worse fit than the interactive model, this would be an indication of interactions, providing further support for the results of the experiments.

In the additive model, each level of an experimental manipulation was represented by a constant that is added to the trust level at the first measurement. Specifically, the after-interaction measurements were modelled as a sum of the before-interaction measurement in the same experimental condition, a fixed effect μ for each level of the manipulated factors Consistency (k) and

Face Validity (l), and again some random measurement error (F) that is uncorrelated across experimental conditions:

$$T_{ijkl,\ after} = T_{ij,\ before} + \mu_k + \mu_l + \varphi_{jkl} * F_{ijkl}, \qquad \text{where } F \sim N(0, 1).$$

The additive model as described above is not identified. Whereas the absolute effects of the two Endorsement manipulations and the effect of the condition in which process feedback was not available (Study 1) could be estimated without problems, the remaining manipulations of Consistency and Face Validity required one of the effects to be fixed. It was arbitrarily chosen to set the absolute effect of the Low Face Validity manipulation to 0.

The interactive model, in contrast to the additive model, allowed different effects for each of the 12 different experimental conditions. Thus, specific combinations of the three experimental factors may result in specific levels of trust; no additivity is assumed, except that the after-interaction measurements are based on the before-interaction level of trust. The modelling therefore only differed regarding the after-interaction measurements:

$$T_{ijkl,\ after} = T_{ij,\ before} + \mu_{jkl} + \varphi_{jkl} * F_{ijkl}, \qquad \text{where } F \sim N(0, 1).$$

These linear mixed models were fitted to the data using Mx (Neale, Boker, Xie, & Maes, 2003), a general program that estimates model parameters by maximizing the log-likelihood of the raw data. Model comparison is done using a likelihood-ratio test.

Results and discussion.

The additive model proved to fit significantly worse than the interactive model, χ^2 (7) = 55.5, $p < .01$. This indicates that the additive model does not fit the data, and that Endorsement Cue, Consistency and Face Validity indeed interact.

Figure 6 displays the observed means, and the means as expected under the interactive and additive models.

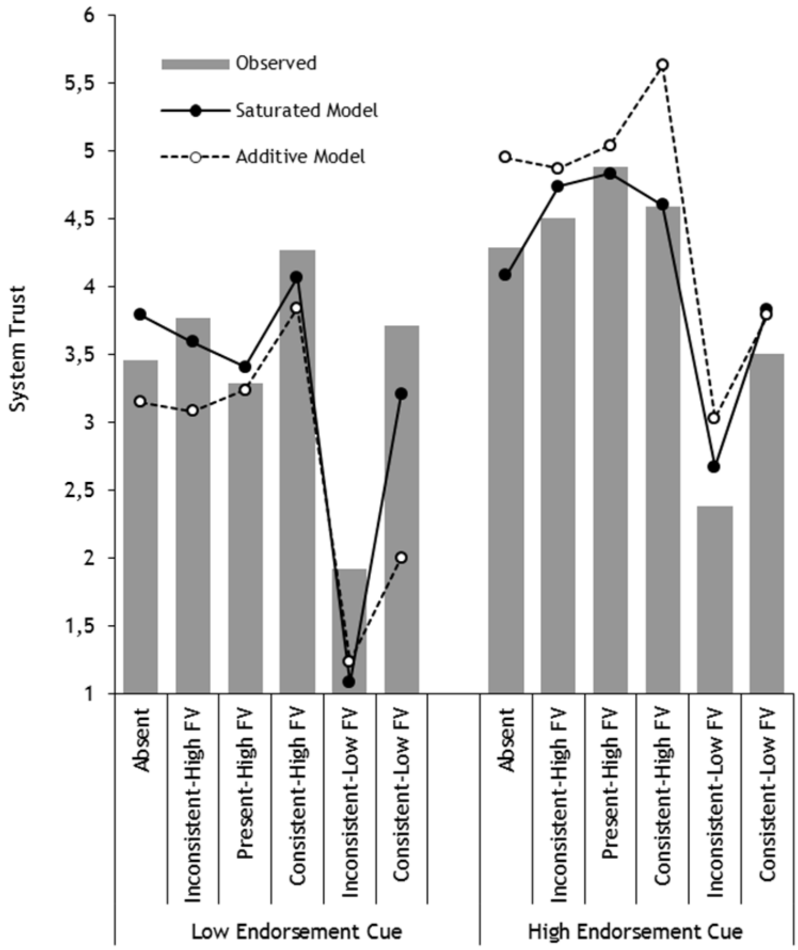

Figure 6. Average observed ratings of system trust, taken after interaction, as a function of Endorsement Cue, Consistency, and Face Validity (FV), compared to ratings expected by the interactive and additive models

Based on the additive model, the effect of Endorsement on trust before-interaction turned out significant; the mean difference between the low and high endorsement cue conditions was M = 1.79 (95% confidence interval: 1.56, 2.02), with higher trust after a high endorsement cue. Absent Process feedback (Study 1) did not result in changes in trust after interaction, the effect being M = -0.22 (95% confidence interval: -0.82, 0.32). Available Process Feedback proved to have the same effect on trust as Inconsistent Process Feedback, the difference between these two conditions being only M = 0.17, and non-significant (95% confidence interval: -0.30, 0.64).

The difference between Consistent and Inconsistent Process Feedback was significant, M = 0.76 (95% confidence interval: 0.37, 1.14); trust was higher after Consistent than after Inconsistent Process Feedback. Compared to Available Process Feedback, Consistent Process Feedback resulted in higher trust, M = 0.59 (95% confidence interval: 0.13, 1.05). The difference between High and Low Face Validity was M = 1.84 (95% confidence interval: 1.44, 2.24).

In summary, Available Process Feedback did not differ from Inconsistent Process Feedback. This supports the assumption made in the discussion of Study 1 about the nature of the displayed routes. These were argued to be somewhat random, causing participants to rely on the Endorsement information in order to interpret what they saw on the screen. Moreover, this overall analysis provided further support for the observation that Consistent Process Feedback resulted in higher trust than Inconsistent Process Feedback. In addition, Process Feedback with High Face Validity also instilled more trust in participants than Process Feedback with Low Face Validity. The additive effect of Face Validity turned out to be far stronger than that of Consistency; the difference between High and Low Face Validity was far greater than that between Consistent and Inconsistent Process Feedback.

However, care should be taken in interpreting these main effects, since the additive model fitted the data significantly worse than the interactive model. This indicates that the different manipulations indeed interacted with each other, as was concluded from the results of the individual experiments. In line with the findings of Studies 2 and 3, the interactive model suggests that the effect of Endorsement on the after-interaction trust levels depends on whether process feedback was Consistent or Inconsistent (see Figure 6). Whereas the additive and interactive model largely agree on the difference in trust between high and low endorsement cue conditions when process feedback was Random, the interactive model estimates this difference to be far smaller in the Consistent Process Feedback conditions. The additive model, however, does not expect a difference in the Endorsement effect between Consistent and Inconsistent Process Feedback. In addition, the interactive model aligns with the results of Study 3, suggesting that the effect of Face Validity depends on the consistency of process feedback; Figure 6 shows that in the Inconsistent conditions the difference in trust between High and Low Face Validity was greater than in Consistent conditions, in line with what was found in Study 3.

General Discussion

The reported studies present a number of interesting phenomena pertaining to situations in which users of systems have different kinds of information at their disposal. Study 1 showed that cue effectiveness could be moderated by the availability of process information. When process feedback was absent, the influence of the cue, given beforehand, diminished over time; when process feedback was present, trust ratings remained fairly stable over time. Study 2 showed that process feedback with a random appearance caused a similar pattern of trust ratings as the process feedback. Presumably, inconsistent process feedback provided rather indefinite visual information that required the content of the endorsement manipulation for interpretation. Hence, the effect of endorsement information was present both before and after the interaction stage. Nevertheless, randomised routes also resulted in a general decrease of trust ratings. Presumably, randomness was taken as a sign of a system′s inadequacy. Consistent process feedback, on the other hand, apparently did provide information on which trust judgements could subsequently be based; this effect was strong enough to completely annihilate the effects of endorsement information.

Study 3 provided further support for the conclusions drawn in Study 2. Again, a main effect of process feedback consistency was found; trust was shown to be higher after consistent rather than inconsistent process feedback. In addition, the effect of consistency became apparent, for example, from the interaction with face validity: the effect of the latter factor was stronger when process feedback was randomised. This could be interpreted as an indication of the "informative content" of consistent process feedback. Process feedback being consistent may have reduced the need to rely on other information, i.e. face validity. Randomness would likely have caused participants to put more emphasis on the face validity of the information. This is conformance with the notion that inconsistent process feedback, as opposed to consistent, provides little information that can be used to build trust; specifically, this lack of information apparently caused the face validity of process feedback to gain weight in trust judgements.

The overall analysis (Study 4) provided additional support for these findings. Whereas the additive model provides information about the magnitude of the manipulations' independent effects on trust, the significantly worse fit of this model with the means observed in the three experiments, compared to the interactive model, indicates that these data cannot be explained by mere additive effects of the manipulations. In line with the findings of Studies 2 and 3, the predictions of the interactive model suggest that the effect of the endorsement manipulation on the after-interaction trust levels depends on whether consistent or inconsistent process feedback had been encountered. Similarly, the interactive model also indicates that the effect of face validity depends on the consistency of process feedback, which is congruent with the results of Study 3.

Furthermore, these results suggest that consistency and face validity of process feedback do not necessarily represent two sides of the same coin. Whereas Study 2 shows that process feedback consistency causes trust to increase, Study 3 indicates that agreement with route advice can cause trust to rise independently of consistency. This finding was supported by the results of the overall analysis, which also showed independent, additive effects of both process feedback consistency and face validity, with the size of the latter exceeding that of the former.

The staked credit measure was employed both as a means to increase participants' task commitment and as an alternative trust measure. The willingness to stake credits on the outcome of the automatic route-planning mode implies a willingness to be vulnerable to its actions, and, as such is a reflection of their trust in the system (cf. Berg, Dickhaut, & McCabe, 1995). Nevertheless, the effects of endorsement and process feedback found in the trust measures did not always materialize on the number of credits. As argued before, a possible explanation lies in the nature and timing of the various manipulations. Whereas endorsement cues were administered before interaction with each route planner, it could well have affected the credits staked in the all trials, especially the first few, and the same goes for face validity of process feedback. Process feedback consistency, however, could only be assessed after at least a number of trials, and its effect on later trials may have been compensated by the lack of effect on the first. In light of these considerations, however, it is noteworthy that the average number of staked credit correlated quite reasonably with after-interaction trust ratings in studies 2 and 3. In addition, these two measures showed the same pattern of results in the additional analyses in Study 3, showing that the effects of the independent variables on both could not be explained by similarity and agreement ratings. We therefore regard the staking of credits as a worthwhile addition to both these studies and future research into system trust.

These findings shed new light on the assumption that indirect information is easily overruled by direct information (Yuliver-Gavish, 2011; Arion *et al*., 1994), for instance, because of the latter's higher informational content (Arion *et al*., 1994). Indeed, process feedback, as a type of direct information that is not accompanied by right/wrong outcome feedback, is capable of overruling indirect information. Whether or not direct experiences provide more information than indirect experiences, however, seems to depend in part on the nature of the former. As Study 2 suggests, if process feedback is consistent, it may indeed overrule the less informative indirect information. The manipulation checks of Study 2 revealed that participants, according to expectation, expressed a stronger belief that the system's output was governed by fixed rules when process feedback had been consistent compared to when it had been random. This provides some support for the notion that consistency facilitates the formation of beliefs about the system′s functioning. The manipulation checks of Study 3, however, only showed a mere trend in this direction. Furthermore, these data seem to indicate that in the case of inconsistent feedback indirect information is actually necessary for interpretation; the data of Studies 2 and 3 clearly showed that this particular combination of direct and indirect information resulted in persistence, rather than extinction of the effect of endorsement information. A further indication that inconsistent feedback required users to call upon other information available was found in the greater influence of face validity on trust in the inconsistent conditions.

An alternative explanation would hold that, rather than fostering a sense of understanding, the system’s behaviour should be seen as either conforming or disproving prior expectations. Merritt and Ilgen (2008) showed that users’ trust before interaction, i.c. their propensity to trust, influenced how the system’s objective characteristics were perceived, and that the resultant perceptions of these characteristics influenced post-task trust. They argued that trust prior to interaction created expectations for system performance, and that the effect of performance on after-interaction trust would depend on the correspondence between performance and expectations. Indeed, they found that high performance resulted in high trust ratings when propensity to trust was also high, and vice versa; contrarily, no effect of performance was found when propensity to trust was low. In the studies reported here initial trust, i.e., trust prior to interaction, was not incorporated by measuring trust propensity but rather by manipulating endorsement cues. Nevertheless, our results could be seen as congruous with Merritt and Ilgen's in underscoring the importance of perceptions of behaviour, rather than objective behaviour, and the interaction of behaviour with initial trust.

Our results differ from Merritt and Ilgen's (2008), however, in the nature of the interaction. If randomised, as opposed to consistent, process feedback is assumed to be a sign of a system′s inadequacy, and the results of studies 2 and 3 support this assumption, then their explanation would imply that consistent process feedback after a high endorsement cue would result in the highest trust levels, and randomised process feedback after a low endorsement cue in the lowest. The after-interaction trust levels in studies 2 and 3, however, deviate from this pattern. For instance, in Study 2, low endorsement with consistent process feedback resulted in ratings similar to high endorsement and consistent process feedback. In Study 3, the combination of high endorsement and consistent process feedback did not result in trust ratings exceeding those in other conditions, regardless of face validity. In addition, the number of staked credits shows that when process feedback was consistent, neither congruence nor incongruence with expectations created by endorsement cues mattered much, if anything. Process feedback either confirming of contradicting expectations can therefore not sufficiently explain these results.

Additionally, one could argue that the interaction of endorsement cues with process feedback could be explained by a low endorsement cue evoking a certain alertness, causing participants to watch system behaviour more closely. Thus, manipulations of consistent versus inconsistent process feedback may only affect trust levels when preceded by a low endorsement cue, as seems to be the case in the after-interaction trust measures in Study 2. Indeed, there is ample evidence to suggest that negative information weighs more heavily in decisions and evaluations (for instance, see Rozin & Royzman, 2001). However, if a minority cue, as opposed to a majority cue, would indeed cause alertness, this would also have resulted in interactions of consensus with process feedback on the checks in studies 2 and 3. However, such interactions could not be shown for any of these checks. Moreover, the pattern of results found in Study 3 are as much the result of changes in the consistent as in the inconsistent process feedback conditions, after both low and high endorsement cues, and as such contradict this alternative explanation.

The concept of process feedback, and its possible beneficial effects on a user's understanding of system functioning has received only scant attention, since it has found its way into theories on system trust a few decades ago (e.g., see Lee & Moray, 1992; 1988). Although Dzindolet *et al*. (2003) tested the effects of understanding of a system′s processes on trust, the emergence of understanding from actually observing process feedback remained obscured. Therefore, these experiments represent a first attempt at uncovering how process feedback plays a role in the development of system trust, via understanding.

For practitioners in the field of user-system interaction, user-system design or software engineering these findings have a number of ramifications. For instance, our results suggest that it is important to not only develop software and hardware that performs as it should, but also to take into account how their advice to the user comes about, and to communicate this to users – especially for novice users. Moreover, we would like to point out that findings such as these, rather than complicate practitioners' work, offer new possibilities to optimise user-system interaction. Specifically, the notion that direct experience with systems has more facets than previously imagined causes the calibration of system trust to no longer be a mere matter of designing for perfect user advice, but also of designing for transparency (cf. De Visser *et al*. 2014; Helldin *et al*., 2013; Ososky *et al*., 2014; Thill, Hemeren, & Nilsson, 2014). Everything a system "does" may cause users to wonder about what happens inside the black box that technology often is, and factors as consistency and face validity may offer new ways to guide this process, and thus novice users' trust in the optimal direction.

Much work needs to be done to fully uncover all trust-relevant aspects of direct information. Since interactions with systems normally entail both outcome as well as process feedback, it seems worthwhile to study not only the former, which yields feedback in clear right/wrong verdicts, but also the latter, that may contain far subtler clues and is subject to interpretation. To fully understand users' perceptions of, and interactions with systems, what matters is not just what these systems do, but also how they do it.

Key points:

- Direct experience with technology is generally considered important for building users' trust, but research is limited to the effects of system output, i.e. the system either yielding accurate of correct solutions or not.
- In the absence of direct experience, novice users are assumed to base their trust solely on indirect information, such as the opinions of other users.
- Research reported here shows that direct experience may also be obtained from interacting with the system, i.e. a route planner, even though concrete, verifiable system output (i.e., right or wrong routing advice) is absent.
- Consistency in process feedback, i.e., the routes suggested by the system, may enable users to make inferences about its underlying processes, and thus increase trust.
- Consistent, as opposed to random, process feedback overrules the effect of indirect information on trust of the user in the system.

References

Allen, R. B. (1997). Mental models and human models. In M.Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of human-computer interaction* (pp. 49-63). Elsevier Science.

Arion, M., Numan, J. H., Pitariu, H., & Jorna, R. (1994). Placing trust in human-computer interaction. In *Proc. 7th European Conference on Cognitive Ergonomics (ECCE 7)* (pp. 353-365).

Berg, J., Dickhaut, J., & Mccabe, K., 1995, Trust, reciprocity, and social history. *Games and Economic Behaviour, 10*, 122-142.

Briggs, P., Burford, B., De Angeli, A., & Lynch, P. (2002). Trust in online advice. *Social Sciences Computer Review*, *20*, 321-332.

Carroll, J. M. & Olson, J. R. (1988). Mental models in human-computer interaction. In M.Helander (Ed.), *Handbook of human-computer interaction* (pp. 45-65). Elsevier Science Publishers.

De Visser, E. J., Cohen, M., Freedy, A., & Parasuraman, R. (2014). A design methodology for trust cue calibration in cognitive agents. In R. Shumaker & S. Lackey (Eds.), *Virtual, augmented and mixed reality. Designing and developing virtual and augmented environments* (Vol. 8525 Lecture Notes in Computer Science, pp. 251-262). Heidelberg: Springer.

De Vries, P., Midden, C., & Bouwhuis, D. (2003). The effects of errors on system trust, self-confidence, and the allocation of control in route planning. *International Journal of Human-Computer Studies*, *58*, 719-735.

De Vries, P. & Midden, C. (2008). The effect of indirect information on system trust and control allocation. *Behaviour and Information Technology,* 27(1), 17-29.

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, *58*, 697-718.

Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., & Treinen, M. (2001). What makes web sites credible? A report on a large quantitative study. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY,USA, March 2001.ACM Press.* 61-68.

Fogg, B. J. & Tseng, H. (1999). Elements of computer credibility. *Proceedings Conference on Human Factors in Computing Systems* 80-87.

Giddens, A. (1990). *The consequences of modernity*. Stanford, CA: Stanford University Press.

Halpin, S., Johnson, E., & Thornberry, J. (1973). Cognitive reliability in manned systems. *IEEE Transactions on Reliability*, *R-22 (3)*, 165-169.

Helldin, Tove, Falkman, Göran, Riveiro, Maria, Dahlbom, Anders, & Lebram, Mikael. (2013). Transparency of military threat evaluation through visualizing uncertainty and system rationale. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics. Applications and services* (Vol. 8020, Lecture Notes in Artificial Intelligence, pp. 263-272). Heidelberg: Springer.

Jensen, C., Davis, J., & Farnham, S. (2002). Finding others online: Reputation systems for social online spaces. *Proceedings Conference on Human Factors in Computing Systems*, *4*, 447-454.

Lee, J. & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*, 1243-1270.

Lee, J. D. & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, *40*, 153-184.

Lee, J. D., & See, S. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46* (1), 50-80.

Lerch, F. & Prietula, M. (1989). How do we trust machine advice? In G.Salvendy & M. Smith (Eds.), *Designing and using human-computer interfaces and knowledge-based systems* (pp. 410-419). Amsterdam: Elsevier Science.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, *13*, 334-359.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review*, *23*, 473-490.

Merritt, S. M. (2011). Affective processes in human-automation interactions. *Human Factors, 53*(4), 356-370.

Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors, 50*(2), 194-210. doi: 10.1518/001872008X288574

Muir, B. M. (1988). Trust between humans and machines, and the design of decision aids. In E.Hollnagel, G. Mancini, & D. D. Woods (Eds.), *Cognitive Engineering in Complex Dynamic Worlds* (pp. 71-84). London: Academic Press.

Muir, B. M. (1989). *Operators' trust in and use of automatic controllers in a supervisory process control task*. Doctoral thesis, University of Toronto, Canada.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling*. (6th edition ed.) Virginia Commonwealth University, Box 900126, Richmond, VA 23298: Department of Psychiatry.

Ososky, Scott, Sanders, Tracy, Jentsch, Florian, Hancock, Peter, & Chen, Jessie YC. (2014). *Determinants of system transparency and its influence on trust in and reliance on unmanned robotic systems.* Paper presented at the SPIE 9084, Unmanned Systems Technology XVI, June 3, Baltimore, Maryland, USA.

Regan, D. T. & Fazio, R. H. (1977). On the consistency between attitudes and behavior: Look to the method of attitude formation. *Journal of Experimental Social Psychology*, *13*, 28-45.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, *49 (1)*, 95-112.

Riley, V. (1996). Operator reliance on automation: Theory and data. In R.Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications* (pp. 19-36). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Rozin, P. & Royzman, E. B. (2001). Negativity bias, negativity dominance and contagion. *Personality and Social Psychology Review*, *5*, 296-320.

Sebrechts, M. M., Marsh, R. L., & Furstenburg, C. T. (1987). Integrative modeling: Changes in mental models during learning. In W.Zachary & S. Robertson (Eds.), *Cognition, Computing and Cooperation* (pp. 338-398). Norwood, NJ: Ablex.

Sheridan, T. B. & Hennessy, R. T. (1984). *Research and modeling of supervisory control behavior*. Washington, DC: National Academy Press.

Thill, S., Hemeren, P. E., & Nilsson, M. (2014). *The apparent intelligence of a system as a factor in situation awareness*. Paper presented at the IEEE International Inter-Disciplinary Conference on

Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), March 3-6, San Antonio, TX, USA.

Verberne, Frank M. F., Ham, Jaap, & Midden, Cees J. H. (2012). Trust in smart systems: Sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. *Human Factors, 54*(5), 799-810. doi: 10.1177/0018720812443825

Woods, D. D., Roth, E. M., & Bennett, K. B. (1987). Exploration in joint human-machine cognitive systems. In W.Zachary & S. Robertson (Eds.), *Cognition, Computing and Cooperation* (pp. 123-158). Norwood, NJ: Ablex.

Yuviler-Gavish, N. and D. Gopher (2011). Effect of descriptive information and experience on automation reliance. *Human Factors, 53*(3), 230-244.

Zuboff, S. (1988). In the Age of the Smart Machine: The Future of Work and Power. New York: Basic Books.

Short author biographies

Dr. Peter W. de Vries is assistant professor at the department of Psychology of Conflict, Risk, and Safety, University of Twente, The Netherlands. He previously worked at Eindhoven University of Technology, The Netherlands, where he received his PhD in 2005, in the field of human-technology interaction.

Dr. Stéphanie van den Berg is associate professor at the department of Research Methodology, Measurement and Data Analysis, University of Twente, The Netherlands. In 2003 she received her PhD in the field of psychology, at Eindhoven University of Technology, The Netherlands.

Dr. Cees Midden is professor of psychology since 1991 and works at the department of Human Technology Interaction, School of Innovation Sciences at Eindhoven University of Technology, The Netherlands.