Is Waze joking? Perceived Irrationality dynamics in user-robot interactions.

Amélie Abadie
ESCP-Europe
amelie.abadie@escpeurope.eu

Samuel Fosso-Wamba
Toulouse Business School
s.fosso-wamba@tbs-education.fr

Kevin Carillo
Toulouse Business School
k.carillo@tbs-education.fr

Olivier Badot
ESCP-Europe
mailto:obadot@escpeurope.eu

Abstract

Artificial intelligence, as well as the use of smarter and smarter systems are gradually pervading our everyday life. While AI-embedded systems are increasingly becoming 'human-like', individuals tend to fictionize interpersonal relationships with them. We conceptualize the notion of Perceived Irrationality as the discrepancy between an individual's expectations and his/her perceptions towards a smart system's recommendation. We then develop a conceptual model that aims at better understanding the inner mechanisms that govern perceived irrationality. This research opens up a vast uncharted research territory that proposes to adopt a long-term relational lens towards the study of humans / smart systems interactions.

1. Introduction

Artificial intelligence is gradually transforming businesses by changing the way they interact with customers and deliver products and services [1]. For instance, mobile device users increasingly rely on smart assistants such as Siri, Alexa, or Google Assistant, to get information, to purchase goods, or else be delivered tailormade services. However, despite major technological improvements in the domain of artificial intelligence, it is rather frequent that Siri or Alexa provide wrong answers which raises customers' frustration. Similarly, when using the AI-enabled GPS system Waze in your car, it may tell you to go left, where a huge sign says "Danger!". You then start believing that Waze may not be the best buddy for your trips. As Waze suggests us directions through travel, A-I embedded apps in smartphones or appliances recommend suitable options along our daily path. Many examples show how these robots, that is to say intelligent agents, may rather frequently act irrationally or at least look irrational in the eye of the user. The rationale behind such statement is that technology is and will be inherently imperfect for the simple reason that it is designed by irrational agents, human beings. In short, technology often fails (e.g. a GPS signal can be lost and all software code contains bugs).

AI-embedded smart devices and IT artifacts keep developing in terms of their capacity to interact, adapt and tailor-make personal exigencies. In addition, if assistants are mainly dedicated to operating when asked, they now are increasingly initiating the conversation thank to improving push notifications, reminders and maintain a ''partial presence'' [63]. The development of relationships between "thinking' devices and humans has never been this usual and promises a long-term phenomenon that may redefine the nature and structure of our society and the role of human beings in our world. Consumer robotics has taken up 7.1% of the whole IT market [56] and are predicted to bring 10 million US$ in 2022 in addition to the 103 million US$ predicted for personal automated vehicles [57]. The growing cohabitation of machines and persons makes room for consideration of complementarities [17]. According to neo-classical economics, rationality embodies the belief that agents pursue without boundaries the best solution that optimizes their isolated interests. However, the idea of rationality seems far away from human cognition but rather reflective of algorithmic intelligence [19, 24]. In other words, person-smart system relationships could be seen as the assembly of the rationality of logics and the irrationality of psychics. Moreover, today digital natives attribute more feelings, beliefs and values to products and services in a world where social coercion is everywhere [6]. Digital transformation also changes paradigms. For instance, researchers started to pair notions of personal traits with intelligent agents [47]. In this paper, we argue that the same prevails, with AI-embedded IT devices (that we commonly call personal robots) but even to a greater extent: individuals

URI: https://hdl.handle.net/10125/59938
ISBN: 978-0-9981331-2-6

attribute feeling, beliefs, and values to intelligent machines and then consequently tend to fictionize interpersonal relationships with the personal smart devices with which they often interact.

This research develops the concept of perceived irrationality from past literature on human rationality and Human-Machine Interaction. We then provide a conceptual model that aims at answering the following research question:

What is the impact of perceived irrationality on user-intelligent system interactions and relationship building mechanisms?

The model is built on the assumption that users fictionize interpersonal relationships with smart agents (at least with the ones they regularly interact with such as smart agents) and that irrational recommendations often occur and have an important impact on their synthetic kinship.

Next section investigates the theoretical foundations surrounding the notion of human and intelligent system irrationality to shape the concept of perceived irrationality. We develop a model that aims at explaining the perceived irrationality dynamics in user-Artificial Intelligence interactions along with its impact on their relationships. Finally, we discuss potential contributions and future directions in a conclusive part.

2. Literature review

2.1. Robots and Humans as friends and foes

There is common agreement about the overall rationality of Information Technology based on the acknowledged assumption that mathematics optimize outcomes in a world where determinism prevails [19, 38]. Artificial intelligence is embedded into intelligent agents: The robot is presently conceived as an Intelligent Agent (1) extracting contextual information and (2) possessing the features to (3) act in this environment. Contextual Data joins the idea of embodiment of robot behavior mentioned in You and Robert [64] and the notion of reaction links to Russel and Norwig [48]. In addition, we follow Dautenhahn [16] cognitive characterization of a robot as ‘’a machine that makes decisions on its own and solves problems’’ of a user. Waze, Siri are enveloped into sensitive hardware from which they capture contingent data to offer a responsive action through Artificial Intelligence. We though consider Smart Apps as intelligent agents, namely robots. The essence of Artificial Intelligence seems to be infused with rationality: their role is to optimize actions according to given situations encountered by users [48]. AI-embedded IT devices tend to be consulted in a repeated and frequent manner: the user requests and the smart device fulfils the user’s wishes. Consequently, reciprocity, an important aspect of interpersonal trust, surrounds many machine-to-human interactions [4, 61]. In other words, a user transposes human-to-human aspects to non-interpersonal interactions with a smart system. It can be argued that users are likely to project human-like criteria such as honesty, or sanity on the agent performance [39]. They should have social skills [16, 64]. For instance, people tolerate better machine failure whenever the failure is communicated in a polite manner [23]. Moreover, the spread of digital private assistants enlarges identities from physical boundaries to digital-selves [6]. Intelligent devices thus allow self-extension and self-expansion for individuals [31] and reveals high accuracy in predicting profound aspirations [17, 18, 19]. We can conceptualize a human-intelligent agent relationship as a seemingly interpersonal relationship created in the user’s mind where one is the intuitive one and the other is the pragmatic successful counterpart [4, 34].

2.2. Human versus smart-devices irrationality

Research from several disciplines has highlighted the overall irrational nature of individuals [50, 59, 62]. Past research has, for instance, worked on the irrational nature of individuals as consumers, emphasizing the fact that trade is rather driven by “Animal Spirits” [2, 33] than considerate arrangements. Besides, users, managers, and investors overall tend to satisfy desires in an imperfect manner [51, 62] rather than conscientiously implementing strategies aiming at maximizing potential outcomes. Meanwhile, agents evaluate potential outcomes according to limited contextual resources and capacities [38, 52]. They also face cognitive limitations: aversion for losses, preference for status-quo, time inconsistency or relative value of gains according to wealth [32]. Humans forget, choose without full consideration with past experience and imperfectly accumulate information [20, 41]. In the opposite direction, myopia and preferences for hedonic or instant outcomes rather than long-term plans [45, 54], tend to enlarge the gap between present decisions and future stakes [36, 32, 45]. Emotions also play an important part in trade or purchasing behaviors through notions such as confirmation bias, social influence [62], success memory or beliefs [59]. Moreover, the so-called “Sunk Cost fallacy” is another example of the emotional nature of human beings [59]. It states that agents care for past investments in present decisions and tend to support plans that did not turn out to be profitable. In market finance, research has shown that traders are not mathematically optimizing portfolio but rather listening to their intuition, values or to stories [2]. In 1979, John Elster [22] distinguishes rationality from optimization of Nature. Emphasizing the dynamic aspects of choice and its scale, norms and willpower parameters, the author summaries holistically in 4 essays the present references. Comparison between human and animal reign lead toward a philosophical definition of rationality. One may be surprised to hear that animals are often better optimizers than humans. People tend to intuitively believe that irrationality and emotions are rooted into instinct. Yet, it has been shown that primary behaviors of animals are highly rational [23].

Irrationality does not pertain to human nor animals. Smart systems and Information Technology, to some extent, are infused with irrationality. Since technology is designed by human beings who are inherently irrational, then technology is imperfect in nature: technology fails! A first piece of illustration lies in the core principle that governs free/open source software development: all software code contains bugs, which justifies the necessity to 'open' the source code and make it freely accessible, sharable and modifiable. As Eric Raymond puts it "given enough eyeballs, all bugs are shallow" [48]. Other sources of 'imperfection' in the specific case of AI-embedded devices may concern the dysfunction of some of the inbuilt censors or any data-generating devices, engendering erroneous recommendations from algorithms due to wrong/incorrect data (in terms of not accurately representing a certain reality). In a similar vein, smart machines often rely on machine learning algorithms (such as deep learning) to generate evolving sets of rules from which recommendations are generated. If datasets are not large enough or else contain erroneous data, then "thinking" machines have no choice but to act irrationally, that is to say by not providing optimal recommendations or not performing optimal actions. Even though they show better impartiality [34, 40], intelligent devices may fail to function [48, 14]. They also have difficulties to understand new environments (because they always look for reference points) [55] and to communicate [13].

Furthermore, another source of irrationality from the user side may pertain to the interests from the agent (often the company) that designs personal smart systems [21, 27]. In certain cases, like recommendation agents such as the one designed by Amazon, intelligent recommenders can be profit-driven [27]. Such smart robots then appear perfectly rational in the eye of the sellers or designers, but may be perceived as irrational to the customers that do not seize the instrumental logic behind certain actions or recommendations.

Nevertheless, irrationality may appear to be a matter of adjustment to humanism [37]. Certain thinkers argue that the gap between robotics and societies is bound to shorten since transhumanism and the humanization of technologies will eventually meet [6]. For instance, AI users tend to request more and more reciprocity, kindness and flexibility from personal devices [37, 39] but also more entertainment and 'foolishness' [20]. Consumers frustrations and requests increasingly shape the robots into affective creatures [37, 44], confidents [18] and even "lovotic" sweathearts [49]. Vinciarelli et al deepen the tie between Technology and Psycholgy and emphasize the need to improve technological identification of users' weak signals [61].

Table 1 summarizes the different aspects that reflect the irrational nature of human beings and that were identified in this research.

Table 1. Human irrationalities.

Concept	Definition	Theories / References
Limited capacity to evaluate present stakes	Difficulties to appreciate clearly all potential opportunities and information.	Search costs [53, 54], information asymmetries [2, 26] or limited attention [38], confirmation bias (Shafir, 1993), status-quo [32]. relative gains [32]
Limited capacity to evaluate past stakes	Backward time inconsistency, difficulty to appreciate clearly all accumulated information.	Bounded memory [41], success memory, trust [59] Sunk Cost fallacy [31]
Limited capacity to evaluate future stakes	Forward time inconsistency: difficulty to appreciate long term outcomes and information.	Butterfly effect [36] overconfidence [2], Myopia, preference for instant outcomes [32].
Limited capacity to maximize outcome	Tendency to choose suboptimal options even in full awareness of past, present or future stakes.	Emotional choice, Social choice [32], Loss aversion [32], stories [2], hedonism, beliefs [59, 62].

3. Theoretical Developments and Research Model

3.1. Perceived Irrationality

In this paper, we argue that the notion of perceived irrationality of intelligent agents stands from a user's perception of incongruity, incomprehension, or a perceived failure during the human-machine interaction [12]. We define perceived irrationality as the difference between a user's ex-ante mental appraisal of a given task or situation [3] and the perception of the actual recommendation provided by the smart system. For instance, when an AI-embedded GPS system is asked to provide an itinerary between point A and B at a given time and under given conditions, the user will make a mental effort to estimate and characterize the itinerary and will then compare it to the one generated by the system. The more the discrepancy between the intuitive appraisal of the itinerary and the actual itinerary computed, the more the user will perceive the recommendation to be irrational (See Figure 1 below).

$$\textbf{Perceived irrationality}_i^j = \Delta(\textbf{Mental Apraisal}_i - \textbf{Perception}_i^j)$$

For a user i to a AI j.

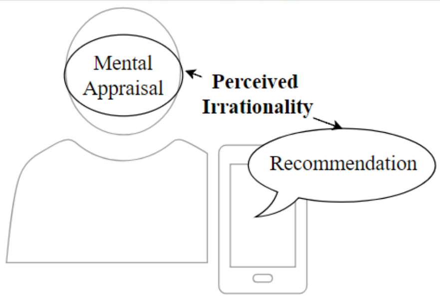

Figure 1. Perceived irrationality as the distance from appraisal to perceived signal.

3.2. Research Model

Table 2. Concepts and definitions

Concept	Definition	Mental appraisal	Perception	Perceived irrationality
Perceived irrationality	Difference between the user ex-ante appraisal [3] of smart device action and the perception of the effective recommendation.			
Perceived control	The perception by the user that she/he is in control of the interactions with a smart system and can somehow influence recommendations (inspired from [1]).	+	=	
Perceived transparency	The ability of a smart device, perceived by the user, to signal accessible and understandable information and extract information from the user [58].	=	+	-
Perceived consequences	Relative, anticipated weight of pursued recommendation into the subject's life.	-	+	--
Perceived complexity	Refers to the perceived behavioural complexity of a request/task [60]. It is the difficulty to predict and control the results of a complex task.	-	=	-
Affective trust	Trust based on feelings of protection, emotions and perception of reciprocity [30].	=	+	-
Cognitive trust	Trust associated to a cognitive estimation of future performance and reliability from past knowledge [30].	=	=	=
Post-recommendation evaluation	Formal valuation of past performance and reliability at time T.	=	+	-
Intention to follow recommendation	The intention of a user to act according to the recommendation provided by a smart system.			

By arguing that users build interpersonal relationships with AI-embedded devices with which they often interact, we develop a conceptual model (see Figure 2) that emphasizes the importance of cognitive and affective trust (key notions in interpersonal relationship building mechanisms) and that explains the user/intelligent system relationship interplay in the specific context of irrational recommendations. The model identifies factors that impact perceived irrationality by influencing one of the two components, or both, that define perceived irrationality (mental appraisal or recommendation evaluation). Perceived irrationality is triggered when a system provides counterintuitive outcomes. As a puzzling failure in its a priori essence, a smart machine considered as acting irrationally would lead to anger, perplexity and mistrust toward the technology or perhaps toward the body (company, brand) that designed the system [23, 25]. It may also bring a user to reconsider his/her trusting belief towards the intelligent agent and may result in the user disregarding the provided recommendation. The following proposition 1 reflects this statement (also presented in Figure 2 below)

Proposition 1: Perceived irrationality has a negative effect on the intention to follow the recommendation.

Perceived irrationality is grounded in an experience of active communication with AI-embedded systems. Even though the user feels fully in control of his/her anticipation and his/her perception, many situational parameters influence both user's opinion and smart device effective advice [13]. The utility attributed to a smart agent is often directly linked to its function [10, 48]. We infer that the more complex the smart device's function is perceived by the user, the more tolerant he or she will be tolerant to perceived irrationality. Our main idea behind this statement is that a user viewing a more complex task will have lower expectations in terms of performance [43]. His/her anticipated action is more likely to be close to the perception of effective action in case of perceived irrationality. Perceived complexity is attributed to a device, we therefore place this concept on the right side in Figure 2 which represents Human-Machine confrontation in the case of perceived irrationality. Our reflection leads to the below proposition 2.

Proposition 2: Perceived complexity has a negative effect on perceived irrationality.

Research on purchasing decisions in the marketing literature has shown that a user observes a given proposition, decides whether to follow it or not and eventually evaluates the situation ex post [9]. Individuals tend to consider the process of following a recommender as an economic lottery [39]. He or she anticipates gains/losses and accounts for risks. We define perceived consequence as the perceived gain or loss resulting from following the smart system advice. As agents are loss-averse [31], the present research refines the concept of consequence as the relative magnitude of pursued recommendation in the subjects life. Therefore, it goes beyond the question of confirmation or refutation of a desire and embodies long and short-term upheavals [53]. We infer a higher stake increases the intention to follow the recommendation. The larger are perceived risks, the lower is one self-confidence and the higher is reliance to the smart device. Responsibility is increasingly devoted to the smart agent as anticipated consequence expends. We summarize this in proposition 3.

Proposition 3: Perceived consequences of following the recommendation has a positive impact on the intention to follow the said recommendation.

AI-embedded devices may be designed with features and options that allow users to manipulate an intelligent system's advices [5]. The perception of control towards a given intelligent machine may increase perceived irrationality as it prevents a user from closing the gap between anticipation and perception. Meanwhile, a user may decide to follow a recommendation that is perceived as irrational by thinking that post-corrections of the recommendation will be possible. Proposition 4 and 5 articulate these subtleties.

Proposition 4: Perceived control has a positive effect on perceived irrationality.

Proposition 5: Perceived control has a negative effect on the intention to follow recommendation.

Data is disclosed by the user and extracted from smart device sensors. It reduces information asymmetry between the two parties [24]. We can assimilate perceived irrationality to the notion of bias in a contract between two agents. As the agency cost in Contract Theory increases with asymmetry of information [26]. We infer that the gap between anticipation and reception shrinks when there is information transparency between the two parties. On the other hand, users have the tendency to look for reasons when facing product failures [23]. Therefore, the more the smart agent communicates information to the user, the easier is the attribution of logical foundations to perceived irrationality. Shared information grants more credibility to recommendation in irrational actions. In addition, users react more charitably when they can communicate their disappointment directly to the machine, instead of dealing with engineers [58]. The ability of an AI-embedded system to react to a user's feelings may give the user a perception of reciprocity [43, 61] and responsibility [28]. We thus argue that the information disclosed between a user and a smart system may mitigate the inconsistency between a user's expectations and a smart agent's recommendation, but also reduces the perception of irrationality towards a given recommendation through the introduction of cooperation and credibility. Proposition 6 theorizes the impact of transparency.

Proposition 6: Transparency has a negative impact on perceived irrationality.

Because the Internet is exposed to privacy spillovers, the lack of sensitive proximity tends to be very adverse [42]. The notion of trust is central in e-service adoption and use [25]. Research has shown that trust shall rather be seen as a dynamic process rather that a static state. For instance, selling and forgetting are considered as already ancient strategies in online e-services [15]. Aktinson et al. claim [4] that "trust in automation cannot be thought of a state, nor can it be of as a single construction or continuum", but rather as an interpersonal dynamic process. Besides, since user-smart device interactions tend to have similarities with the development of interpersonal relationships, we argue that trust is an important factor that characterize user-smart system relationships [43]. There may also exists long-term mistrust [30]. Instant perceived irrationality may have lower effects on user trust while iterated long-term foolishness may hinder the use of the device. We distinguish the effects of cognitive and affective trust on perceived irrationality [30, 53]. Affective trust is based on the user's opinion from past experiences with the smart agent and on beliefs from the user's peers. In addition, affective trust is based on feelings of protection, reputation, emotions and perception of reciprocity [30, 37]. Therefore, we think affective that trust will distort the ex-ante perspective of the user. The user would be inclined to higher tolerance on the intelligent system as a friend is caring about a friend [49]. Proposition 7 synthesizes our consideration about the influence of affective trust.

Proposition 7: Affective trust has a negative effect on perceived irrationality by the user.

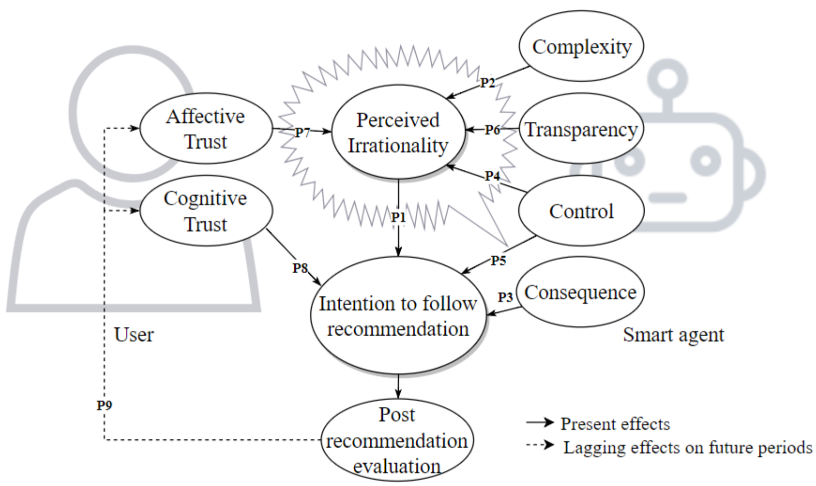

Figure 2. Model of perceived irrationality in user-smart system interaction

Affective trust is independent from instant cognitive trust. Even though a friend is acting poorly, affective trust remains. Cognitive trust is grounded into perceptions of performance and reliability from past knowledge [30]. Cognitive trust does not influence the perception of irrationality but does influence the intention to follow a recommendation that is perceived as irrational. Indeed, we argue that a user feeling a high cognitive trust towards a smart system based on past interactions, will follow the smart device's recommendation and disregard his/her own perception of irrationality. In other words, the user will think that the intelligent machine tends to be always right based on previous positive experiences, and despite the current perception of irrationality. We illustrate this in Figure 2 as a user attribute in the dynamics of perceived smart agent irrationality.

Proposition 8: Cognitive trust has a positive effect on the intention to follow the recommendation.

Finally, in line with the expectation-confirmation theory [9], we develop the dynamic effect of perceived irrationality on future irrationality perceptions towards a smart system. A positive post-evaluation of a smart device's recommendation having been perceived as irrational, will increase trust in future interactions with this intelligent machine. Consequently, a positive post-recommendation evaluation shall impact positively both affective and cognitive trust [8]. We hypothesize that positive post evaluation will enhance perceived intelligence and relevance of Artificial Intelligence embedded actions. Therefore, the consumer will gain confidence in her ability to personalize and build affection on a non-emotional machine while building objective trust in its performance through experience. Therefore, proposition 9 states the dynamic loop between trust and evaluations.

Proposition 9: Performance evaluation has a positive effect on present affective and cognitive trust.

4. Potential contributions and conclusion

The present paper provides a wide range of future directions for research. We extend this research aspiration through this section around perceived irrationality and our conclusive remarks.

We develop the notion of perceived irrationality, as both a confrontation between expectation and confirmation [8, 9] and confrontation between animal and robotic spirits. Artificial intelligence may appear irrational in the eye of the user based on the perceived inconsistency between what the user thinks and what he sees. We identify factors that may have an impact on perceived irrationality and thus affecting human-machine interactions [4]. In addition, we identify potential lagging dynamics from trust theories [30] which applicability is hypothesized to extent to the user-smart system interaction context. The example of mistrust in autonomous cars is pregnant. As customers anticipate loosed control, producers communicate experiments with elegant cars to show transparency and to increase cognitive and affective trust in customers. They also stress the intelligence behind technologies to show that complexity and responsibility is rigorously handled, hoping for the benevolence of consumers.

This research in progress conceptualizes perceived irrationality as a gap between human expectation and smart device perceived actions. Its decomposition (see Figure 1) raises questions about its causes: how do we distinguish personal expectations from the perceptions of effective failure? It could help legislate resolutions of conflicts in after sale innovation and service: is the human or the artificial intelligence guilty of a mistake? Is human testimony always objective about machines?

Because machines came from industrials to shoppers, smart agents and human tend nowadays to bond into intimate relationships. As smart systems have increasing implication in our daily habits, research tends to consider them as intimate companions [4]. Therefore, future research could turn to Relationship Science. Is there a level where trust transforms into actual dependence? As opposed, could perceived irrationality trigger dangerous mistrust of a consumer? Examination of irrationality perception in technology-dependent subjects could deepen knowledge about trust.

A relevant research focus should aim at measuring the impact of trust on perceived irrationality: how long does it takes for a consumer to trust enough her devices so to forgive it when she perceives irrationality? Could trust transform perceived irrationality into perceived treason?

In addition, high contributions could emerge from the study of perceived irrationality on surrounding non-user humans. How peer's control affects perceived irrationality? Are there differences in interaction when we alleviate the assumption of privacy? How can social interaction with other humans influence the irrationality perceived from an intelligent system? Further research along this avenue could broaden the status of smart robot from exclusive confidents to comparable members of a crowd. It would open Artificial Intelligence research to strategic reactions such as jealousy or manipulation. As technology is widely used in public space and embodies a social dimension, it could help retailers and e-businesses balancing between high personalization and social collaboration.

Transparency is key to long, dynamic friendships. It possesses an important weight on perceptions [9] and reactions to disappointment. Shared knowledge among communities promotes service and fantasies [11]. The question of confidentiality remains: should perceived irrationality be transparent to the user only or to other users, to increase Eric Raymond's [48] collaborative performance? Could information overload trigger perceived irrationality?

The preceding interrogations about legislation, information or peer credibility over smart agents raise a concern about the pervasiveness of robots. More than issues of performance, the perception of irrationality into robot raises ethical concerns [35]: when does this interpersonal relationship imagined by the user violate privacy?

Many researchers anchored the subject of interpersonal relationships into customer relations, brand content or product performance [25]. Some applied it to e-commerce, looking at how recommendation systems influence individuals in their daily habits or purchases [17, 10, 11]. In addition, researchers have seriously started contemplating the collaboration between men and smart devices as an interpersonal collaboration [4] and as an emphasis of human skills [29]. User-smart system interactions is increasingly relevant as smart AI-embedded products prosper in the market while the traditional frontier between impartial Artificial Intelligence and emotional human blurs [7, 29, 72, 44]. Past research in Human Robot Interaction regards conflict between Human personality and Robot features [64] and considered them as personality traits [47]. It also underlines the need for intelligence in identifying social signals [64]. When theory closes the statics between Humans and Machines, this research extends it to a dynamic framework by considering perceived irrationality as an event into reactions rather than ex ante human-robot settlements.

This framework applies to smart agents which are regularly used, as only interactions could create perceived irrationality. Initial usefulness in daily habits appears to be key in initiating an interpersonal fiction. However, as the interpersonal relationship takes its roots into the user's mind, we think the lack of initiative of a robot affects little the dynamics of perceived irrationality. In addition, as smart assistants gain autonomy, they increasingly inform us without our request [67]. This research is still at a very early stage. Details about the methodology that will be followed is still currently under debate. The next step of this research may consist of empirically testing the conceptual model through quasi-experiments. A first argument in favor of such research method is our consideration for emotional and affective processes. These aspects require a near observant of reactions. Also, since we explore the dynamic frame of user-smart system relationships, we aim at studying a longitudinal panel of users to whom we would provide devices. As irrationality may arise from the user perception and may have no link with our intentions, we would request regular feedbacks from the user to identify emerging perceived irrationalities. Self-selection bias may be a key issue in autogenerated perceived irrationalities. We would, after a period necessary for interdependence and appreciation of technology, randomize a planned irrationality from the systems. Randomization would also discriminate cognitive from affective reactions.

The ideal frame of experiment would involve a new (non-personal) intelligent system to observe the potential development of a relationship. The use of Collaborative Robots, safe and tractable intelligent agents initially used by industrials, is taken into consideration. Indeed, they enable perceived and effective irrationality without harming the experimenter. We plan to observe collaboration overtime through a range of tasks asked to cases. We aim at implementing randomization through critical incidents

We are also considering the implementation of a simulation research design as this would provide complementary insights. We would first simulate a model between two fully rational agents. We would then identify how deviations from this simulated norm could affect the iterated process of interaction. Our definition of perceived irrationality opens to both human and intelligent machine failures to understand the situation. To disentangle perception from rational observation, the idea of comparing autonomous appliances to users using smart assistants for a same task is also considered. The first case offers an almost fully rational interaction while the second alleviates the rationality of the protagonists. Irrationality could therefore rise from planned programming error in both cases and lead to difference-in-differences analysis. We aim at studying how irrational humans respond to rational smart agents and vice versa, and how two irrational agents may influence interactions through various scenarios.

Irrationality in Artificial Intelligence appears as a failure. However, it is rather humanism and creates room for thought in fields formerly restricted to myths.

5. References

[1] R. Agarwal and E. Karahanna, "Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage," MIS Quarterly, vol. 24, no. 4, 2000, pp. 665-694.

[2] G.A. Akerlof, G. A., and R.J. Shiller, Animal spirits: How human psychology drives the economy, and why it matters for global capitalism., Princeton University Press, Princeton NJ USA, 2009.

[3] M. B. Arnold, Emotion and personality, Columbia University Press, New York NY USA, 1960.

[4] D. Atkinson, P. Hancock, R. Hoffman, J.D. Lee, E. Rovira, C. Stokes C. and A. Wagner, "Trust in computers and robot: the uses and boundaries of the analogy of interpersonal trust.", Proceedings of the Human Factors and Ergonomics Society, 56th Annual Meeting, Boston MA, 2012.

[5] C. Y. Baldwin and K. B. Clark, "Modularity in the Design of Complex Engineering Systems," in D. Braha, A. A. Minai, and Y. Bar-Yam, Complex Engineered Systems: Science Meets Technology, Springer Berlin Heidelberg, Berlin, 2006.

[6] R. W. Belk, "Extended Self in a Digital World," Journal of Consumer Research, vol. 40, no. 3, 2013, pp. 477-500.

[7] R. Belk, "Anthropomorphism and Anthropocentrism", Association for Consumer Research, vol. 44, pp. 42-47, 2016.

[8] R. Bhattacharya, T. M. Devinney, and M. M. Pillutla, "A Formal Model of Trust Based on Outcomes," The Academy of Management Review, vol. 23, no. 3, 1998, pp. 459-472.

[9] A. Bhattacherjee, "Understanding Information Systems Continuance: An Expectation-Confirmation Model," MIS Quarterly, vol. 25, no. 3, 2001, pp. 351-370.

[10] I. Benbasat and B. Xiao, "E-commerce product recommendation agents: use, characteristics, and impact," MIS Quarterly, vol. 31, no. 1, 2007, pp. 137-209, 2007.

[11] Y. Bouzid and R. Vanheems, "Comportement web-to-store : vers une nouvelle logique de contrôle de soi ?," Management & Avenir, vol. 71, no. 5, 2014, pp. 189-200.

[12] J. Caelen and H. Nguyen, "Traitement des incompréhensions et des malentendus en dialogue Homme-Machine. " , Actes de la 13ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'2006), Leuven, 2006.

[13] D. Cameron, J.Aitken, E. Collins et al. "Framing Factors: The Importance of Context and the Individual in Understanding Trust in Human-Robot Interaction.", International Conference on Intelligent Robots and Systems (IROS) 2015, Hamburg, 2015.

[14] J. Carlson, R. R. Murphy and A. Nelson, "Follow-up analysis of mobile robot failures", Proceedings. ICRA '04. 2004 IEEE International Conference on Robotics and Automation, New Orleans LA, 2004.

[15] J. Mou and J.F. Cohen, "Trust and online consumer health service success: A longitudinal study," Information Development, vol. 33, no. 2, 2017, pp. 169-189.

[16] K. Dautenhahn "Socially intelligent robots: dimensions of human–robot interaction", Phil. Trans. R. Soc. B, vol. 362, 2007, pp. 679-704.

[17] T.H. Davenport and J.G. Harris, "What people want (and how to predict it).", MIT Sloan Management Review, vol. 50, no.2, , 2009, pp. 23-31.

[18] S.S. Davidowitz, Everybody lies: Big Data New Data and what the Internet can tell us about who we really are, HarperCollins, New-York NY USA, 2017.

[19] R. Dawes, "The Robust Beauty of Improper Linear Models." American Psychologist, vol. 34, 1979, p. 571-582.

[20] M. Dodgson, D. M. Gann, and N. Phillips, "Organizational Learning and the Technology of Foolishness: The Case of Virtual Worlds at IBM," Organization Science, vol. 24, no. 5, 2013, pp. 1358-1376.

[21] G. Ellison and F. Ellison Sara, "Search, Obfuscation, and Price Elasticities on the Internet," Econometrica, vol. 77, no. 2, 2009, pp. 427-452.

[22] J. Elster, Ulysses and the Sirens: Studies in Rationality and Irrationality, Cambridge University Press, Cambridge MA USA, 1979.

[23] V. S. Folkes and B. Kotsos, "Buyers' and sellers' explanations for product failure: Who done it?," Journal of Marketing, vol. 50, no. 2, 1986, pp. 74-80.

[24] J. R. Galbraith, "Organization Design: An Information Processing View," Interfaces, vol. 4, no. 3, 1974, pp. 28-36, 1974.

[25] D. Gefen, E. Karahanna, and D. W. Straub, "Trust and TAM in online shopping: an integrated model," MIS Quarterly, vol. 27, no. 1, 2003, pp. 51-90.

[26] S. J. Grossman and O. D. Hart, "An Analysis of the Principal-Agent Problem," Econometrica, vol. 51, no. 1, 1983, pp. 7-45.

[27] A. Hagiu and B. Jullien, "Why do intermediaries divert search?," RAND Journal of Economics, vol. 42, no. 2, 2011, pp. 337-362.

[28] F.A. Hayek, "The use of knowledge in society", The American Economic Review, vol. 35, no. 4, 1945, pp. 519-530.

[29] D. Hoffman, T. Novak T. and H. Kang, "Anthropomorphism From Self-Extension and Self-Expansion Processes: an Assemblage Theory Approach to Interactions Between Consumers and Smart Devices", Association for Consumer Research, 44, 2016, pp. 42-47.

[30] D. Johnson and K. Grayson, "Cognitive and affective trust in service relationships," Journal of Business Research, vol. 58, no. 4, 2005, pp. 500-507.

[31] D. Kahneman, "Maps of Bounded Rationality: Psychology for Behavioral Economics," American Economic Review, vol. 93, no. 5, 2003, pp. 1449-1475.

[32] D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," Econometrica, vol. 47, no. 2, 1979, pp. 263-91.

[33] J.M. Keynes, The General Theory of Employment, Interest and Money, Macmillan, London, 1936.

[34] J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan, "Human Decisions and Machine Predictions" The Quarterly Journal of Economics, vol. 133, no. 1, 2018, pp. 237-293.

[35] P. Lin , K. Abney., and G. A. Bekey, Robot Ethics: The Ethical and Social Implications of Robotics. The MIT Press, Cambridge MA, 2014.

[36] E. N. Lorenz, "Deterministic Nonperiodic Flow," Journal of the Atmospheric Sciences, vol. 20, no. 2, 1963, pp. 130-141.

[37] S. Magids, A. Zorfas and D. Leemon D.,“The new science of customer emotions.”, Harvard Business Review, vol. 93, no. 11, 2016, pp. 66-68.

[38] Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay, "Revealed Attention," American Economic Review, vol. 102, no. 5, 2012, pp. 2183-2205.

[39] J. E. McEneaney, "Agency Effects in Human–Computer Interaction", International Journal of Human–Computer Interaction, vol. 29, no. 12, 2013, pp. 798-813.

[40] P. Meehl, Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence. J. Aronson, Lanham MD, 1954.

[41] S. Mullainathan, "A Memory-Based Model of Bounded Rationality" The Quarterly Journal of Economics, vol. 117, no. 3, 2002, pp. 735-774.

[42] E. Pantano, B. Nguyen, C. Dennis and S. Gerlach, Internet Retailing and Future Perspectives. 2nd Edition., Routledge, New-York NY USA, 2017.

[43] A. H. Peter, R. B. Deborah, E. S. Kristin, Y. C. C. Jessie, J. d. V. Ewart, and P. Raja, "A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction" Human Factors, vol. 53, no. 5, 2011, pp. 517-527.

[44] R. W. Picard, Affective computing. MIT Press, Cambridge MA, 1997.

[45] J. Quiggin, "A theory of anticipated utility," Journal of Economic Behavior & Organization, vol. 3, no. 4, 1982, pp. 323-343.

[46] E. Raymond, "The cathedral and the bazaar," Knowledge, Technology & Policy, vol. 12, no. 3, 1999, pp. 23-49.

[47] L. P. Robert, "Personality in the Human Robot Interaction Literature: A Review and Brief Critique", Proceedings of the 24th Americas Conference on Information Systems, (AMCIS 2018), New Orleans LA, 2018.

[48] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach. Pearson Education, London, 2003.

[49] H. A. Samani, A. D. Cheok, F. W. Ngiap,, A. Nagpal, and M. Qiu, "Towards a formulation of love in human-robot interaction.”, Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN),Viareggio, 2010.

[50] P. A. Samuelson, "A Note on the Pure Theory of Consumer's Behaviour," Economica, vol. 5, no. 17, 1938., pp. 61-71

[51] E. Shafir, I. Simonson, and A. Tversky, "Reason-based choice," Cognition, vol. 49, no. 1, 1993, pp. 11-36.

[52] H. A. Simon, "Rational choice and the structure of the environment," Psychological Review, vol. 63, no. 2, 1956, pp. 129-138.

[53] H. A. Simon, "Motivational and emotional controls of cognition," Psychological Review, vol. 74, no. 1, 1967, pp. 29-39.

[54] G. E. Smith, M. P. Venkatraman, and R. R. Dholakia, "Diagnosing the search cost effect: Waiting time and the moderating impact of prior category knowledge," Journal of Economic Psychology, vol. 20, no. 3, 1999, pp. 285-314.

[55] T. Sotiropoulos, H. Waeselynck, J. Guiochet, and F. Ingrand, "Can Robot Navigation Bugs Be Found in Simulation? An Exploratory Study," in 2017 IEEE International Conference on Software Quality, Reliability and Security (QRS), Prague, 2017.

[56] Statista DC, “Robotics market distribution worldwide in 2016, by use case.“, 2017. [https://www-statista-com.revproxy.escpeurope.eu/statistics/661893/worldwide-robotics-market-distribution-by-use-case/]

[57] Statista Tractica, ”Projected non-industrial robotics market size worldwide in 2022, by major segment (in million U.S. dollars”, 2017. [https://www-statista-com.revproxy.escpeurope.eu/statistics/760223/worldwide-non-industrial-robotics-market-revenue-by-segment/]

[58] S. S. Sundar and C. Nass, "Source Orientation in Human-Computer Interaction: Programmer, Networker, or Independent Social Actor," Communication Research, vol. 27, no. 6, 2000, pp. 683-703.

[59] J. Tirole, "Rational irrationality: Some economics of self-management," European Economic Review, vol. 46, no. 4-5, pp. 633-655, 2002.

[60] A. Tiwana, Platform Ecosystems, Aligning Architecture, Governance, and Strategy, 1st Edition, Morgan Kaufmann, Burlington MA, 2014.

[61] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: Survey of an emerging domain" Image and Vision Computing, vol.27, no.12, 2009, pp.1743–1759.

[62] J. Waldfogel, "Does Consumer Irrationality Trump Consumer Sovereignty?," The Review of Economics and Statistics, vol. 87, no. 4, 2005, pp. 691-696.

[63] I. Warren, A. Meads, S. Srirama, T. Weerasinghe, and C. Paniagua, "Push notification mechanisms for pervasive smartphone applications. " IEEE Pervasive Computing, vol. 2, pp. 61-71, 2014.

[64] S. You and L. P. Robert, "Emotional Attachment, Performance, and Viability in Teams Collaborating with Embodied Physical Action (EPA) Robots", Journal of the Association for Information Systems, (JAIS), vol. 19, no. 5, 2018