Stereotype Content Model (SCM) and Chatbots / Conversational Interfaces – an experiment comparing trust, competence and warmth dimensions

Roger Seiler
Zurich University of Applied Sciences (ZHAW)
Steffen Müller
ZHAW School of Management and Law
Markus Beinert
Hochschule Weihenstephan-Triesdorf

Cite as:
Seiler Roger, Müller Steffen, Beinert Markus (2019), Stereotype Content Model (SCM) and Chatbots / Conversational Interfaces – an experiment comparing trust, competence and warmth dimensions. *Proceedings of the European Marketing Academy*, 48th, (8468)

Paper presented at the 48th Annual EMAC Conference, Hamburg, May 24-27, 2019.

Stereotype Content Model (SCM) and Chatbots / Conversational Interfaces – an experiment comparing trust, competence and warmth dimensions

Abstract

With the rising popularity of chatbots this research paper seeks to answer the question if the stereotype content model (SCM) applies to the domain of chatbots or broader to conversational interfaces. This study answers this research question by conducting an experiment containing different stereotypes (lovable star and incompetent jerk). The SCM applies to the domain of chatbots. The lovable star stereotype chatbot is perceived as more trustworthy as well as more competent and warmer compared to the incompetent jerk. Participants pointed out that they want to know if they are talking to a chatbot and not to a real human. Nevertheless, further research is required regarding traditional text chatbots because the lovable star did not show higher trustworthiness than the text chatbot. Furthermore, these research results suggest, that data privacy is a further, important aspect as customers typically share information when engaging in a conversation with a chatbot.

Keywords: chatbots, stereotype content model (SCM), trust

Track: Digital Marketing & Social Media

1. Introduction

Chatbots are taking over today's technology industry (Dale, 2016, p. 811). Within a year over 100’000 chatbots were released on Facebook’s Messenger platform (Johnson, 2017). Google’s Assistant, Apple’s Siri, Amazon’s Alexa and Microsoft’s Cortana are examples of modern digital assistants that understand natural spoken language and answer questions asked by users. Bots can automate simple tasks such as booking hotel rooms, or complex tasks like, health and insurance advice or online shopping guidance. The term «Robo Advisor» has emerged for chatbots with an advisory character such as legal or financial advice. These developments have led to predictions about their impact in the near future. Foye (2017) estimates that $8 billion in cost savings per year are to be expected by 2022 in the financial and insurance sector. Gartner estimates that by 2020 chatbots may have taken over 25 percent of customer service communication (Gartner, 2018) and according to Business Insider 80 percent of businesses will have implemented a form of chatbot by 2020 (Business Insider, 2016). The chatbot topic seems to bear high potential and is thus relevant. Nevertheless, there is evidence that suggests that this does not hold true for user acceptance. 80 percent of users seeking healthcare of financial advice state that they would prefer talking to a human rather than to a chatbot, even though 90 percent of users are aware and believe that companies are ready to make use of this technology (Hopping, 2018). Lack of trust and data privacy concerns are potential explanations (Følstad, Nordheim, & Bjørkli, 2018; Hopping, 2018).

The stereotype content model (SCM) proposes that humans judge other humans in social interaction using the two dimensions likability (warmth) and competence (Fiske, Cuddy, Glick, & Xu, 2002). Highly competent and warm people are referred to “lovable stars”, whereas incompetent and “cold” people are referred to as “incompetent jerks” (Casciaro & Lobo, 2005). These dimensions or traits affect trust. Therefore, we ask the research question how and if the SCM can be applied to the context of chatbots. This is relevant because we argue that this could help companies use the potential of chatbots and bridge the trust gap that currently seems to be existing.

2. Theoretical background

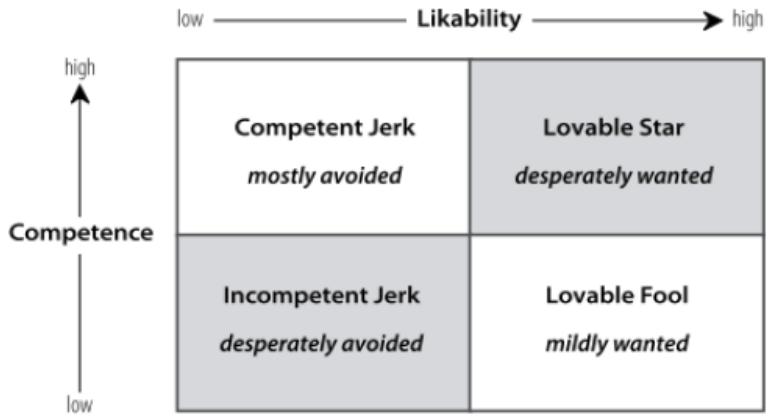


Figure 1: stereotype content model of (Casciaro & Lobo, 2005)

Previous studies have shown that warmth and competence have a positive effect on trust (Seiler, Hari, & Kavci, 2015, 2016). Warmth and competence induce trust (Aaker, Garbinsky, & Vohs, 2012). Not only do high warmth and high competence lead to admiration of a brand, it also leads to a higher willingness to purchase since the brand is perceived more trustworthy (Aaker et al., 2012). However, warmth has a larger impact on trust than competence (Aaker et al., 2012). This is confirmed in a sales context where warmth is the primary dimension (Arndt, Evans, Landry, Mady, & Pongpatipat, 2014) and inlin with other studies identifying warmth as a main driver of preferences towards products or countries (H. Xu, Leung, & Yan, 2013). Moreover, the trust model by Martin and Martin (2014), displays that high affinity and high competence lead to higher trust(Martin & Martin, 2014). One can therefore argue that companies should aim to build chatbots high in warmth and competence, in order to be perceived as trustworthy and positively influence conversion rates.

First impressions instantly form when perceiving a counterpart in a face-to-face or mediated communication setting. A person needs less than 100 milliseconds to make a judgement about the counterpart being perceived (Willis & Todorov, 2006). This effect applies the context of salespeople too. A salesperson is judged within seconds and clients decide, whether the salesperson is helpful or too pushy (Ambady, Krabbenhoft, & Hogan, 2006). In the context of the SCM, warmth and competence judgments are made within the first few seconds of an encounter with another person (Fiske, Cuddy, & Glick, 2007).

In the context of mediated communication e.g. chatbots, social presence plays an important role when choosing to interact with computer agents (K. Xu & Lombard, 2017). Social presence can be described as a psychological state where virtual actors are perceived as

real actors (Lee, 2004). Lombard and Ditton (1997) identified two types of social presence: a virtual character presenting social cues within a medium and the medium itself presented as social actor (Lombard & Ditton, 1997). When the medium signals social cues, people are inclined to see the medium as a real person (K. Xu & Lombard, 2017).

Social presence of embodied avatars shown on company websites significantly influence the perceived trust and emotional appeal of the website, if the avatars are perceived as pleasant (Etemad-Sajadi & Ghachem, 2015). Social presence is a main driver of trust and online purchase intentions (Lu, Fan, & Zhou, 2016) as well as showing positive effects on perceived trust in recommendation agents (Hess, Fuller, & Campbell, 2009).

A further theoretical frame is provided by CLT. Levels of construal increase (high level) with psychological distance e.g. if a decision is in the far future and decrease (low level) when the psychological distance is closer (Trope & Liberman, 2010). Trope et al. 2007 state that low level construal are concrete whereas high level construal are more abstract. Yan, Sengupta, & Hong (2016) suggest that verbal representations are more abstract and visual more concrete representation and furthermore state that people rely on visual processing for proximal events. This view is shared by Aggarwal & Zhao (2015) stating local perceptual processing for low levels of construal. We therefore argue, that in situations where social presence is high, such as this is the case if a person relates to the picture of a chatbot, the CLT theory can further explain how people acquire information. Furthermore, the interaction with a chatbot is proximal and therefore people should rely more on pictures than this would be the case in a more distance event. Hernandez, Wright, & Ferminiano Rodrigues (2015) conclude that in distance events (high construal) benefit based appeals are more persuasive than attribute-based ones, whereas in low construal attributes are equally persuasive and if low levels are temporarily induced even more persuasive than benefit appeals.

3. Research objective and hypotheses

The goal of this study is to verify if the SCM applies to chatbots and helps to close the gap trust research regarding chatbots. By testing if adding pictures of people to a chatbot leads to higher perceived trustworthiness, warmth and competence levels, this research can give guidance to practitioners on reducing concerns customers currently have.

According to the SCM (Casciaro & Lobo, 2005), evidence from research on social presence and first impressions (Willis & Todorov, 2006), Chatbot 1 (lovable star, see Table 2) should be perceived warmer than Chatbot 2 (incompetent jerk), hence stating our first hypothesis (H1). Not only warmth should be perceived on a higher level for Chatbot 1 than

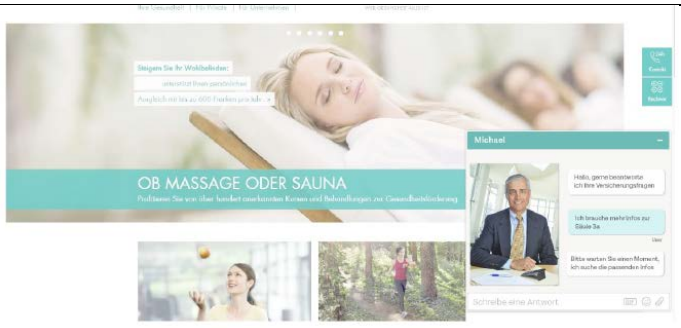
Chatbot 2 but competence too, should be higher for Chatbot 1 than for Chatbot 2, hence our second hypothesis (H2). Compared to the chatbot designed as text only, Chatbot 3, Chatbot 1 should show higher levels of warmth (H3) and higher levels of competence (H4) than Chatbot 3. According to SCM Chatbot 1 should have higher levels of trust than Chatbot 2 (H5) as well as higher levels than Chatbot 3 (H6).

#	Hypothesis	Result
H1	Chatbot 1 (lovable star design) is perceived warmer than Chatbot 2 (incompetent jerk design)	Accepted
H2	Chatbot 1 (lovable star design) is perceived more competent than Chatbot 2 (incompetent jerk design)	Accepted
H3	Chatbot 1 (lovable star design) is perceived warmer than Chatbot 3 (simple text chatbot)	Accepted
H4	Chatbot 1 (lovable star design) is perceived more competent than Chatbot 3 (simple text chatbot)	Accepted
H5	Chatbot 1 (lovable star design) is perceived more trustworthy than Chatbot 2 (incompetent jerk design).	Accepted
H6	Chatbot 1 (lovable star design) is perceived more trustworthy than Chatbot 3 (simple text bot).	Rejected

Table 1: Overview of Hypothesis

4. Methodology

Participants are randomly assigned to one of the groups (between-design) before being redirected to an online survey. A first group of participants interacts with a highly competent and likable chatbot (“lovable star”), a second interacts with an incompetent and unlikable chatbot (“incompetent jerk”), and a third group interacts with a text only chatbot without a picture. The pictures have been validated as stimuli in previous research (Seiler, Hari, & Kavci, 2015, 2016).

Chatbot 1 **“Lovable star”**	

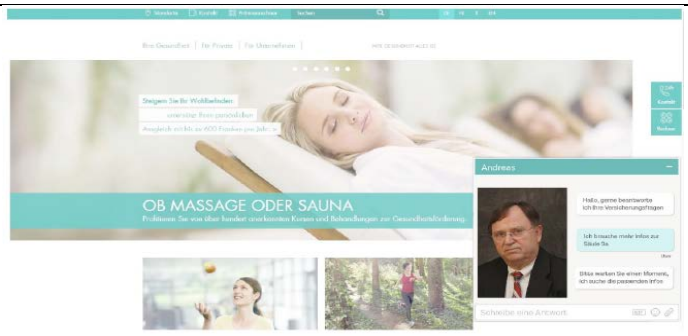
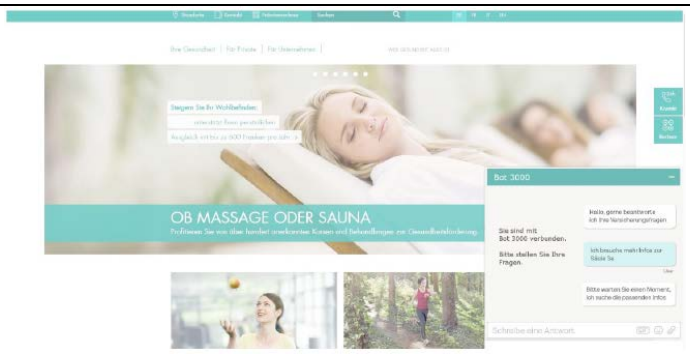
Chatbot 2 **“Incompetent jerk”**	
Chatbot 3 **“Chatbot 3000”**	

Table 2: Chatbots used in this study

The first part of the questionnaire contains demographic and psychographic questions preceding a section with the constructs drawn from research on SCM before asking questions about usage of chatbots as well as information people would be willing to provide to the chatbot. The scales of Reinhard, Messner, & Sporer (2008) are used as these researchers have translated the scales of the SCM to German.

After pretesting the questionnaire, the experiment was conducted in 2018 in Switzerland by sending the online survey’s link to students at the University. In total, 140 respondents completed the questionnaire. The sample is composed of the three groups of Chatbot 1 (n=44), Chatbot 2 (n=50) and Chatbot 3 (n=46).

Before analysing the results, data is cleansed and participants that did not pass the manipulation check, recalling the chatbot’s name, are removed. This leaves 38 subjects in the group of Chatbot 1, 41 in the group of Chatbot 2 and 50 participants in the group of Chatbot 3 as everyone in this group passed the manipulation check. This results in a total of n=129 for further, detailed analysis.

5. Results

Regarding gender, 60 women, 66 men and 3 people not stating the gender compose the sample. The age ranges from 19 to 54 years, with an average of 27 years. All three subsamples do not differ significantly (e.g. gender or age).

Regarding psychographic variables, participants agree the most with regularly using social media (M = 4.94, SD = 1.90), followed by being technology-oriented (M = 4.77, SD = 1.53), being up to date with digital trends (M = 4.03, SD = 1.69), and finally they agree the least with being an early adopter to new technology (M = 3.52, SD = 1.61). The Cronbach's Alpha values of the constructs warmth, competence and trust all pass the minimum requirement of 0.7 (Nunally, 1978) and only in subgroup analysis (e.g. competence Chatbot 1 and trust Chatbot 2) some values are over 0.9 (i.e. .942 and .940) which some researchers argue may indicate redundancy (Tavakol & Dennick, 2011). We therefore conduct hypothesis testing with the constructs without removing any items by using SPSS to calculate t-tests.

Hypothesis testing is conducted by using t-test in order to test for significant differences between the different groups and then deciding whether to accept or reject the hypothesis being tested (an overview is given in Table 1). Warmth scores are significantly higher for Chatbot 1 (M = 4.59, SD = 1.57) than for Chatbot 2 (M = 3.13, SD = 1.390, t (314) = 8.72, $p < .001$). Therefore, hypothesis number one (H1) is accepted. Chatbot 1 (M = 5.01, SD = 1.45) shows higher competence values than Chatbot 2 (M = 4.29, SD = 1.563, t (393) = 4.72, $p < .001$). Thus, hypothesis number two (H2) is accepted. Chatbot 1 (M = 4.59, SD = 1.57) has higher values regarding the warmth construct than Chatbot 3 (M = 3.82, SD = 1.604, t (350) = 4.51, $p < .001$). Therefore, hypothesis number three (H3) is accepted. Levels of pereceived competence are higher for Chatbot 1 (M = 5.01, SD = 1.45) than for Chatbot 3 (M = 3.97, SD = 1.43, t (438) = 7.52, $p < .001$). Hence, hypothesis number four (H4) is accepted. Chatbot 1 (M = 4.51, SD = 1.58) has significantly higher levels of trust than Chatbot 2 (M = 4.15, SD = 1.59, t (393) = 2.25, $p < .05$). Thus, hypothesis number five (H5) is accepted. No significant difference regarding perceived trust of Chatbot 1 (M = 4.51, SD = 1.58) and Chatbot 3 (M = 4.43, SD = 1.30, t (438) = .599, $p > .05$) can be reported. Therefore, hypothesis number six (H6) has to be rejected.

6. Discussion

According to data analysis, the lovable star (Chatbot 1) shows higher levels of warmth, compared to the incompetent jerk (Chatbot 2). The results suggest that levels of competence too, are higher for Chatbot 1 than for Chatbot 2. Therfore, this dimension of the SCM too, applies to the context of chatbots. Trust as well, shows higher levels with Chatbot 1 than 2. Therfore, accordingly this dimension applies to the context of chatbots. Nevertheless, it has to be mentioned that the level of trust of Chatbot 1 did not show significantly higher levels than Chatbot 3. Because the SCM applies to Chatobot 1 and Chatbot 2, both showing pictures and

high social presence, we think the setup of the text only chatbot to hold an explanation. People may have already become accustom to text only chatbots and thus no significant levels of trust are observed. An alternative explanation could be that participants mentioned they would rather interact with actual human beings than with artefacts such as automated chatbots. Therefore, they perceived both Chatbot 1 and Chatbot 3 as being mechanical hence, not showing significant differences. Furthermore, CLT could hold an explanation. The various Psychological distances and construal levels are related respectively influenced by each other (Trope & Liberman, 2010). Therefore, we hypothesise that due to the fact that text only chatbot is less social present higher levels of construal may apply.

Because the level of student participants is high, this sample has to be characterized as student sample. Not all four stereotypes of the SCM were tested, hence results may be different for the other two stereotypes (i.e., “competent jerk” and “loveable fool”). Furthermore, these results may be subject to regional or cultural effects, in this case the context of the greater Zurich Area, and generalization to other regions may not apply.

Research on social presence theory and trust in chatbots may be worthwhile because the lovable star did not show higher levels of trust compared to the text chatbot. If this is due to the experimental design or users being accustomed to text only chatbots through past interactions leading them to feel very comfortable with text only chatbot interfaces may lead to interesting insights. Furthermore, the CLT theory could be further researched and attribute vs. alterantive-based strategies of acquiring information (see Marzocchi, Pizzi, & Scarpi (2016)) may be interesting regarding chatbots.

7. Conclusion

The SCM applies to the context of chatbots. Adding the picture of a lovable star stereotype instead of an incompetent jerk picture of a sales clerk, companies can raise perceived trustworthiness, competence as well as warmth of a chatbot. Past research suggest that higher levels of trust positively influence purchasing intentions (Lu, Fan, & Zhou, 2016). We therefore suggest that pictures in line with the SCM are used instead of any arbitrary picture or avatar in the design of chatbots. The higher levels of trustworthiness have further positive effects on consumer behaviour, such as providing information to the chatbot, which is valuable not only for marketers but the chatbot too, as with more information a better service and better machine learning is to be expected with more and better data. Machine learning could further be an option for delivering segment specific pictures to the chatbot in order to maximise the effects on trust, competence and trust. Furthermore, we suggest A/B-testing

chatbot interfaces as the text chatbot is a valid option as participants may have already been accustomed to this form of interface and therefore changes to known and recognizable interfaces should be executed with great caution, the SCM providing guidance to changes in designing images and avatars of chatbots.

8. References

Aaker, J. L., Garbinsky, E. N., & Vohs, K. D. (2012). Cultivating admiration in brands: Warmth, competence, and landing in the "golden quadrant." *Journal of Consumer Psychology*, *22*(2), 191–194. https://doi.org/10.1016/j.jcps.2011.11.012

Aggarwal, P., & Zhao, M. (2015). Seeing the Big Picture: The Effect of Height on the Level of Construal. *Journal of Marketing Research*, *52*(1), 120–133. https://doi.org/10.1509/jmr.12.0067

Ambady, N., Krabbenhoft, M. A., & Hogan, D. (2006). The 30-Sec Sale: Using Thin-Slice Judgments to Evaluate Sales Effectiveness. *Journal of Consumer Psychology*, *16*(1), 4–13. https://doi.org/10.1207/s15327663jcp1601_2

Arndt, A., Evans, K., Landry, T. D., Mady, S., & Pongpatipat, C. (2014). The impact of salesperson credibility-building statements on later stages of the sales encounter. *Journal of Personal Selling & Sales Management*, *34*(1), 19–32. https://doi.org/10.1080/08853134.2013.870182

Business Insider. (2016, December 14). 80% of businesses want chatbots by 2020 - Business Insider. Retrieved November 24, 2018, from https://www.businessinsider.com/80-of-businesses-want-chatbots-by-2020-2016-12

Casciaro, T., & Lobo, M. S. (2005). Competent jerks, lovable fools, and the formation of social networks. *Harvard Business Review*, *83*(6), 92–9, 149. https://doi.org/10.1037

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, *22*(05), 811–817. https://doi.org/10.1017/S1351324916000243

Etemad-Sajadi, R., & Ghachem, L. (2015). The impact of hedonic and utilitarian value of online avatars on e-service quality. *Computers in Human Behavior*, *52*, 81–86. https://doi.org/10.1016/j.chb.2015.05.048

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What Makes Users Trust a Chatbot for Customer Service? An Exploratory Interview Study. In S. S. Bodrunova (Ed.), *Internet Science* (pp. 194–208). Springer International Publishing.

Foye, L. (2017, July 24). Chatbot Conversations to deliver $8 billion in Cost savings by 2022. Retrieved November 24, 2018, from https://www.juniperresearch.com/analystxpress/july-2017/chatbot-conversations-to-deliver-8bn-cost-saving

Gartner. (2018, February 19). Gartner Says 25 Percent of Customer Service Operations Will Use Virtual Customer Assistants by 2020. Retrieved from https://www.gartner.com/en/newsroom/press-releases/2018-02-19-gartner-says-25-percent-of-customer-service-operations-will-use-virtual-customer-assistants-by-2020

Hernandez, J. M. da C., Wright, S. A., & Ferminiano Rodrigues, F. (2015). Attributes Versus Benefits: The Role of Construal Levels and Appeal Type on the Persuasiveness of Marketing

Messages. *Journal of Advertising*, *44*(3), 243–253. https://doi.org/10.1080/00913367.2014.967425
Hess, T., Fuller, M., & Campbell, D. (2009). Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents. *Journal of the Association for Information Systems*, *10*(12). Retrieved from https://aisel.aisnet.org/jais/vol10/iss12/1
Hopping, C. (2018, February 23). 80% of customers don't trust chatbots for aftersales advice | IT PRO. Retrieved November 24, 2018, from https://www.itpro.co.uk/machine-learning/30606/80-of-customers-dont-trust-chatbots-for-aftersales-advice
Johnson, K. (2017, April 18). Facebook Messenger hits 100,000 bots | VentureBeat. Retrieved November 24, 2018, from https://venturebeat.com/2017/04/18/facebook-messenger-hits-100000-bots/
Lee, K. M. (2004). Presence, Explicated. *Communication Theory*, *14*(1), 27–50. https://doi.org/10.1111/j.1468-2885.2004.tb00302.x
Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, *3*(2), JCMC321-JCMC321. https://doi.org/10.1111/j.1083-6101.1997.tb00072.x
Lu, B., Fan, W., & Zhou, M. (2016). Social presence, trust, and social commerce purchase intention: An empirical research. *Computers in Human Behavior*, *56*, 225–237. https://doi.org/10.1016/j.chb.2015.11.057
Martin, D., & Martin, R. (2014). Towards a model of trust. *Journal of Business Strategy*, *35*(4), 45–51. https://doi.org/10.1108/JBS-05-2014-0053
Marzocchi, G. L., Pizzi, G., & Scarpi, D. (2016). When a picture's worth a thousand words: The effects of visual construal priming on information acquisition and choice. *Marketing Letters*, *27*(3), 487–498. https://doi.org/10.1007/s11002-015-9365-y
Nunally, J. (1978). *Psychometric theory*. New York: McGraw-Hill.
Reinhard, M.-A., Messner, M., & Sporer, S. L. (2008). Explicit Persuasive Intent and Its Impact on Success at Persuasion—The Determining Roles of Attractiveness and Likeableness. *Journal of Consumer Psychology*, *16*(3), 249–259. https://doi.org/10.1207/s15327663jcp1603_7
Seiler, R., Hari, J. J., & Kavci, S. (2015). Changing the warmth and competence dimensions: experimental validation in the context of insurance consultants. In *Academy of Marketing Conference 2015, Limerick, 7-9 July 2015* (pp. 1–9).
Seiler, R., Hari, J., & Kavci, S. (2016). The warmth and competence dimensions: experimental validation in the context of crowdfunding. Presented at the Academy of Marketing - Annual Conference & Doctoral Colloquium, Newcastle.
Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd
Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. https://doi.org/10.1037/a0018963
Willis, J., & Todorov, A. (2006). Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, *17*(7), 592–598.
Xu, H., Leung, A., & Yan, R.-N. (Terry). (2013). It is nice to be important, but it is more important to be nice: Country-of-origin's perceived warmth in product failures. *Journal of Consumer Behaviour*, *12*(4), 285–292. https://doi.org/10.1002/cb.1419
Xu, K., & Lombard, M. (2017). Persuasive computing: Feeling peer pressure from multiple computer agents. *Computers in Human Behavior*, *74*, 152–162. https://doi.org/10.1016/j.chb.2017.04.043
Yan, D., Sengupta, J., & Hong, J. (2016). Why Does Psychological Distance Influence Construal Level? The Role of Processing Mode: Table 1. *Journal of Consumer Research*, *43*(4), 598–613. https://doi.org/10.1093/jcr/ucw045