Metahuman vs. Videoconference: Comparison of Empathy for 3D VR Characters and 2D Human Stimulus

Empathy depends on a person's cognitive processes influenced by quality, presence and flow of the experienced scenario. During the Covid-19 pandemic many personal conversations have been moved from the real world to videoconferencing setups. With the advance of VR technology, possibilities arise to mimic real facial expressions in form of a metahuman character. In this paper we report results from a user study (N=24) comparing the empathy towards a real person in a video call setting and a realistic VR-character designed to mimic facial expressions of the real person. Results show that both conditions evoked the same level of empathy among viewers. The metahuman format showed less variance among the empathy levels. This indicates that VR characters with real facial expressions increase the utility of virtual reality in the realm of interpersonal functioning, which can result in a wide range of application areas such as business meetings, gaming, or healthcare.

CCS Concepts: • **Human-centered computing** → **Interactive systems and tools**; **Collaborative and social computing**.

Additional Key Words and Phrases: matahuman, virtual reality, virtual characters, avatar, empathy

ACM Reference Format:

. 2022. Metahuman vs. Videoconference: Comparison of Empathy for 3D VR Characters and 2D Human Stimulus. 1, 1 (March 2022), 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Communication between people through the support of technologies takes place everywhere and has increased with the Covid-19 pandemic. The possibilities of telecommunication and computer-mediated communication are becoming increasingly present [17]. Beside video conferencing systems such as Zoom or Webex, Virtual Reality (VR) might be a common platform in the future where communication takes place. Thus, this type of virtual interaction is particularly interesting for HCI. Interpersonal characteristics are in focus to make communication in VR as promising as possible. One aspect of interpersonal is a persons ability to be empathetic towards another person and studies have shown that VR has the potential to influence interpersonal emotions such as empathy [18].

Thus, increasing the level of realism for a VR character might also lead to higher empathy. Factors to increase realism are designing the character more human like using real facial expressions taken from a real person and implement it in the VR characters expressions. In order to compare empathy towards such a realistic VR-character with a conventional video call we set up a laboratory study with 24 participants. Its aim was to examine the impact of virtual reality on the enhancement of empathy for another person using realistic avatars with real facial expressions. Thus our research question was: *To which extend does the level of empathy for a realistic VR character differ in comparison to a human stimulus in the form of a conventional video call?*

Author's address:

Manuscript submitted to ACM

In the laboratory study, participants viewed a young female student consulting the student pastoral care because of personal issues. Participants were divided into watching either a metahuman in a virtual reality format or a real human in a two-dimensional format. In the following we present related work on empathy, VR and empathy, as well as virtual characters and empathy. Then we describe the study setup, its results and discuss our findings.

2 RELATED WORK

Empathy has a multi-component characteristic. It can be divided into two different aspects [3]. The affective dimension describe the ability of putting oneself in the perspective of another person including showing appropriate emotional reactions to another person's situation [1, 2]. Cognitive aspects describe the ability of understanding another person's situation and intentions [18]. Davis [3] describes two components that act as antecedents for the feeling of empathy. The first component is the person itself. The persons in the observer position bring characteristics with them that affect their sense of affective and cognitive empathy. The capacity for empathy plays a decisive role. It describes a person's ability to put themselves in a certain position or to have specific reactions to the experience of others. The second component of evoking empathy is the situation. The specific situational context defines what another person experiences. This context determines the degree of empathy of the person observing. The situation influences the observer's reactions to another person, which happens through different dimensions that go hand in hand with the situation. One of these is the emotional intensity of the situation. This strongly influences the reactions of the observer. If a situation involves strong negative emotions, this will evoke strong reactions from the observer. On the other hand, the similarity of the persons involved influences the situation and thus the empathy. The more similar the observer and the target are, the stronger the empathic reaction of the observer and their ability to put themselves in the perspective of the person and understands their situation [3].

Combining empathy with virtual reality opens a new field of application. Studies show that virtual reality experiences can result in a higher level of empathy for shown characters [18]. Various factors influence the perception of empathy in VR. These include presence, quality and flow of the shown scenario [20]. The presence factor is cited as being decisive for empathy in VR. Presence describes the experience of immersion or the state of mind of a person in a virtual environment [9, 14]. It includes the degree to which two people who communicate and interact via a technological medium have the feeling of being together and feeling close [16]. Crucial for the feeling of presence in virtual reality is the given sensory control [14]. Being in virtual reality makes it seem to the observer as if they are close to a person or involved in a situation. This makes it easier to perceive the emotions of the other person and understanding the situation [20]. Empathy and presence also do have common features like the projection of oneself into an environment or the experience of another person [16, 18]. Studies show that the degree of presence in virtual reality is related to empathy. The more realistic a situation in virtual reality is, the better people can empathize [16]. This is because empathic responses are stronger when a person is in the situation. The strength of the situation is thus strongly pronounced. This sense of presence suggests that empathy for another person in VR is promising and may overtake other communication technologies.

When using virtual environments to communicate, a supporting medium to represent oneself virtually is needed. An avatar can represent one's own person as a virtual character. Research on avatars has already shown that, with the appropriate authenticity, they can evoke a similar level of empathy as a comparable human stimulus [17]. To create this authenticity, many details are needed to make the avatar human. Johnson et al. [10] have shown the influence of inauthentic facial expressions. They used computer animation to create such facial expressions and have shown that micro expressions in human emotions are often not correctly reflected. Observers need a long time to get used to the

avatar and to interpret its emotions and behavior correctly due to low recognition accuracy. They found that observers show a mimic response when interacting with another person. The results indicate that if an avatar is in a negative state the observer feels negative emotions too [10]. This concept of matched expressions with another person is called "facial mimicry". It describes the relationship between facial expressions and empathy and is considered crucial for social interaction. [5, 6] If the imitation of facial expressions is not possible, the result is reduced emotion. [8] In order to be able to correctly recognize the emotions of another person and to be empathetic, an avatar should convey correct facial expressions.

Based on this related literature, we can conclude that research on the impact of facial expressions and communication in virtual reality can increase application areas for several HCI-Applications and bring knowledge in the realm of interpersonal functioning. Application areas include empathy training since studies show that virtual reality is a promising tool for improving social skills [11]. This could either be applied in autism to practice social skills and perception taking [4, 11], or to train medical personnel like medical students, physicians or therapists, that show greater competence and better outcomes when having higher levels of empathy [14]. Another application area is transferring empathy in the realm of distant communication with need for empathy like virtual consultation of physicians or psychiatrists or in business to distant experience events or meetings.

3 STUDY DESIGN

The study followed a between subject design. We designed a scenario, which was implemented in a two-dimensional video-setting showing a real human, as well as a metahuman character in VR. Questionnaires were implemented in LimeSurvey[1]. The study was divided into three steps. The first step of the study included a questionnaire containing questions about the participants' demographic data and the Toronto Empathy Questionnaire [21]. Second step was executing either a three-dimensional virtual reality experience or viewing a two-dimensional video call. In the last step, participants were asked to fill in a questionnaire containing questions regarding the empathy for the character and ones own felt similarity to the character, as well as an open-ended question aiming for a self-report regarding the perceived empathy, reasons, presentation and comments.

According to the rules of (*blinded for review*) a formal approval from the ethics board was not required, but we followed our institutions ethics and data protection guidelines. This includes voluntariness, freedom of discontinuation, and deletion of data if requested. No personal data apart from gender, age, prior experience with VR and profession was taken and data was handled completely anonymously. Participants got an information sheet and had to sign an informed consent. Exclusion criteria for recruiting included for example known issues with VR. Participants got no compensation.

3.1 Scenario and Character

The scenario was as follows: Participants received the instruction of being a counselor in student pastoral care. Due to the Covid-19 situation, their institution had to restructure and switched to an online format for meeting students. Participants were being told that they will now see a real student via a digital platform who is consulting the student pastoral care because of personal issues. They were asked to listen carefully and focus on the characters' problems, situation and emotions. The participants of the three-dimensional condition were told that they will see a representative avatar of a real person.

[1] https://www.limesurvey.org

The scenario was self-generated. We scripted a story about the character's living situation, loneliness, overload and life struggles and recorded a student telling the story pretending to be the character with these issues. This allowed us to match the VR-character's visuals to the character in the video call to avoid effects of bias. In addition, this made it possible to show the exact same facial expressions in both conditions. The scenario was tested by fellow students.

For the two-dimensional condition we showed the recorded video in a WebEx Meeting [2] pretending to see the character live. For the three-dimensional condition we created an avatar with MetaHuman[3]. MetaHuman is an Unreal Engine-based framework that allows the creations of high-fidelity digital humans. Allthough there is a variety of different MetaHumans to choose from, for the scope of this study a digital version of the real person in the 2D condition video was created. In order to ensure the visual similiarity between the real person and the MetaHuman, reference pictures were used. The MetaHuman was then exported into the real-time GameEngine UnrealEngine. To record the facial animations an UnrealEngine feature called LiveLink was used. LiveLink is a common interface for streaming and consuming animation data from external sources (for example a mobile app) into the GameEngine. In combination with the iPhones Feature of FaceID, the real-time data of the actor are streamed directly to the animation-ready rigged face of the MetaHuman. Participants in the three-dimensional condition wore the HTC Vive virtual reality system which provided 360° head-tracking and wore a headset providing spacial sound coming from the character. Participants sat on a placed chair during the procedure to ensure everyone having the same distance to toe character. The height of the chair was adjusted so participants viewed the character at eye level.

Fig. 1. Illustration of Video Comparison & Metahuman

3.2 Measurement

To investigate in the evoked empathy for the virtual characters we used different measurements. Since the evoked empathy depends very much on the person, we measured whether the two groups are comparable. For this we needed investigations in the overall capacity for empathy and the extend of similarity. On the other hand, we measured the situational empathy for the character.

To measure the *capacity for empathy* of each participant we used the Toronto Empathy Questionnaire (TEQ) [21]. TEQ is a parsimonious tool to assess empathy through a combination of self-report measures. It consists of 16 items, with each item rated on a 5-point scale, ranging from "never" to "always". For this study, we used TEQ to be able to

[2]https://www.webex.com
[3]https://www.unrealengine.com/en-US/metahuman

compare capacity for empathy for both groups of participants (VR and 2D). For the scoring we summed scores to derive total following TEQ usage instructions.

Table 1. Similarity and empathy items based on [7].

Similarity
I think I have a lot of things in common with Carina, the character. (SI)
I feel similar to Carina, the character. (SI)
Empathy
I can really imagine the thoughts running through Carina's, the characeter's, head. (EM)
I can really feel what Carina, the character, must have been feeling in this situation. (EM)
I can experience the same feelings that Carina, the character, experienced. (EM)
I can take the perspective of Carina, the character, and understand why the situation occurred. (EM)
I can really see myself in Carina's, the characters's, shoes. (EM)
I feel like I can easily take the perspective of Carina, the character. (EM)

To assess participants feeling of *similarity* and *empathy* with the character we adapted items from Haegerich and Bottoms [7] (see Table 1). Adaption was done to fit our scenario (e.g. inclusion of the characters name). Items were rated on a 7-point scale, with a high score indicating strong agreement and a low score indicating strong disagreement. Items were queried after watching the character telling her story. For the scoring we calculated mean values, according to the questionnaire guidelines [7]. Here we have to mention, that it is not possible to clearly attribute meaning to the similarity and empathy scores. Therefore, the scores can only be compared under the same measurement conditions and describe differences in enhanced empathy between groups.

Since all questions are short and concise, we additionally implemented an open self-report to eliminate or exclude possible effects and to get additional answers that open the interpretation of the results. In this context, we added an open question, asking participants for their perceived empathy, reasons, influence of the media and other things they noticed or wanted to mention.

3.3 Participants

In this study, 24 people (9 female, 15 male) between 20 and 36 years (M=25.7, SD=4.29) participated. Participants were recruited among students at (blinded for review) as a convenience sample. Both sub-samples (12 participants for each condition) ran through the same study procedure. Participants in the VR condition were requested to do the study in the lab since the 3D glasses were needed, whereas for the 2D condition were asked to participate online.

4 RESULTS

Data analysis was done in two phases. First, we examined in which extend both groups are comparable. Since empathy is strongly dependent on the person, we wanted to make sure that both groups had about the same capacity for empathy and were equally similar to the character. This ensures that we then in a second step can compare the subsequently examined empathy of both groups towards the character.

In terms of the capacity for empathy, the Levene's test is significant ($p = 0.01$), indicating a violation of the equal variances assumption. Therefore, no t-test for independent samples was performed. There is variance heterogeneity between the two groups 2D (M = 46.58; SD= 4.08) and VR (M=45.33; SD= 6.90). To examine the similarity to character, a

Table 2. Means and standard deviation for the VR and 2D condition

Measure	Virtual Reality (N = 12)		2D Condition (N = 12)	
	Mean	*SD*	Mean	*SD*
Capacity for Empathy	45.33	6.90	46.58	4.08
Extend of Similarity	4.00	1.64	3.58	1.52
Empathy for Character	5.82	0.82	5.51	1.30

Welch's t-test for independent samples was performed. There is no significant difference between the two groups 2D (M = 3.58; SD = 1.52) and VR (M = 4.00; SD = 1.64) in terms of similarity to character (t (21.9) = -0.646, p = 0.525, d = -0.26).

In terms of the empathy for the character, the Levene's test is significant (p = 0.03), indicating a violation of the equal variances assumption. Therefore, no t-test for independent samples was performed. There is variance heterogeneity between the two groups 2D (M = 5.51, SD = 1.30) and VR (M = 5.82, SD = 0.82). Means and standard derivation can be seen in Table 1. Distribution of variances can be seen in Figure 2.

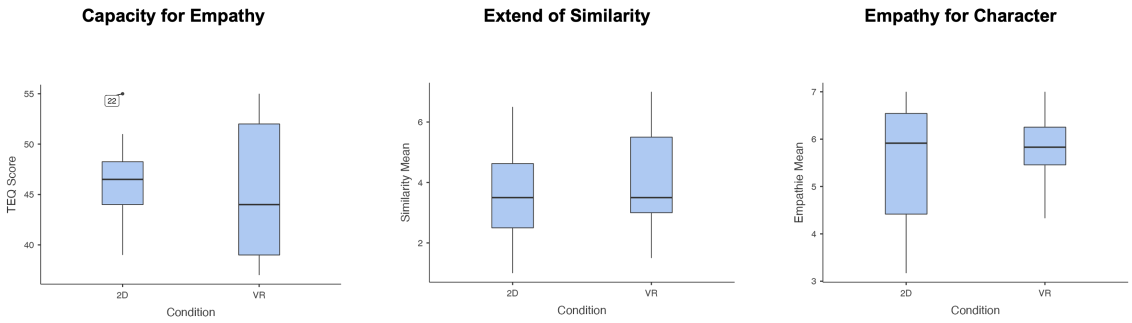

Fig. 2. Distribution of measured values of virtual reality and 2D condition

5 DISCUSSION

Results indicate that comparing the two groups is difficult. They do have equal extend of similarity to the character but differ in the variance of the distribution of values for the capacity of empathy. The 2D group is normally distributed whereas the VR group has a bimodal distribution. This is presumably due to the small sample, which is not representative. The bimodal distribution of the VR condition shows that a part of the participants of this group are very empathic and the other group was way less empathic. To find an explanation for this, we took a closer look at the data and discovered a strong effect between gender and capacity for empathy. Women show significantly higher values than men. However, this cannot be the reason for the bimodal distribution in the VR group, since in relative terms there was one more woman in this group. If we look more closely at the empathy scores of the genders in the two groups, women in the VR group were very empathic and the men are the very opposite. In the 2D group, women were also more empathic on average than the men, but there was less variance, resulting in a normal distribution in the 2D group. There is no obvious explanation for this distribution. We estimate that the sample was chosen unfavorably which is due to the disadvantages of an opportunity sample. Presumably, this effect could have been circumvented by a larger sample.

Nevertheless, the differences in mean value of both groups show nearly similar values for the capacity for empathy. Thus, we were able to perform a pseudo comparison of the groups regarding the empathy for the character. Again, the distributions of the two groups are not statistically comparable. A closer look at the data revealed that the men from the VR group, who indicated a low capacity for empathy are particularly empathetic towards the character and even show higher values here than the women from this group. This raises the question of whether this is due to the sample or whether the visuals have something to do with it. It is also possible that the VR experience can enhance a lot of empathy, especially in less empathic people.

5.1 VR vs. 2D

To answer the research question, the level of empathy for a realistic VR character can be interpreted as similar to a human stimulus in the form of a conventional video call. Overall, mean values of the VR condition were a little higher than the mean values of the 2D condition. In addition, the variance was lower. Thus, participants were in agreement and generally had a high level of empathy for the VR character. For the 2D condition, there was a high variance, and again a bimodal distribution. The data show that most of the women in the group were again very empathetic, although some were significantly less empathic. On the other hand, men in this condition show an equal distribution between the extremes. It would be interesting to find out why men show such a strong variance in empathy in the 2D condition and show strong empathy in the VR condition. This raises the question of whether this is an effect of the VR condition or due to the sample. In conclusion, it can be said that the VR condition not only triggers the same amount of empathy despite being an avatar, but also has a lower variance. Thus, the subjects were less dispersed regarding empathy in the VR condition.

Further insights and possible explanations come with the evaluation of the open-ended question. In general, participants from both groups described that they were very familiar with the scenario and the problems of the character and either had been in a similar situation themselves or had friends and acquaintances who had been in the situation before. All participants were able to empathize with the character through the chosen scenario and the presentation. It is particularly worth mentioning that this was often explicitly mentioned by participants from the VR condition. It is interesting that the participants in the VR condition often mentioned that the facial expressions matched the statements of the character well, which was rarely mentioned by the participants of the 2D condition. Since there were no comments on inauthentic facial expressions, we assume that appropriate facial expressions are taken for granted in a conventional video call and are considered a highlight in VR.

In addition, subjects in the VR condition often mentioned that the conversation seemed very real and authentic to them, as in a real encounter. Regarding the reality of the avatar, the keyword "uncanny valley" was also used twice, which describes an unpleasant impression of a of a humanoid robot that has an almost, but not perfectly, realistic human appearance [19]. In contrast, in the 2D condition, the subjects often mentioned that the situation seemed unreal, since they only had contact via a screen, and that they would have preferred to be in the same room as the character. The video call seemed rather distant. It was also mentioned twice that people are used to video calls through Covid, but that they are still impersonal.

In summary, open answers show that the 2D condition was more often declared as impersonal and distant and the VR condition conveys a real and authentic impression. We assume that the presence and realness created by the virtual reality environment and the avatar with facial expressions have a positive effect on empathy for virtual characters in virtual reality and that these are not only equivalent to video calls but can also surpass them.

5.2 Limitations

Regarding the sample, it should be considered that it was a very small opportunity sample. The subjects are not representative for society. Also, among the subjects were experts in multimedia arts and technology and 3D visualization. Further research could include a bigger and representative sample so that the comparability of groups is given.

Regarding the measurements instruments in this research, the questionnaire to assess the capacity for empathy was very short. A long version could possibly provide better results, but also requires more time to complete. The same applies to similarity to character and empathy for character. We also recommend using psychosomatic measurements to assess empathy. This is recommended by other studies [12, 13, 15] to be a more valid measurement for empathy.

According to the subjects' own statements, the story was familiar to many of them, and it was easy to feel what the character was going through. This may have influenced the results, in that the subjects who had already experienced this were more empathetic and the few who had not yet experienced this were less empathetic. A scenario that is equally known or unknown to all subjects could have provided clearer results.

Concerning the audio, many subjects stated that they paid a lot of attention to it. This may have distracted from the facial expressions. It cannot be excluded from the data that the sound had an influence on the facial expressions. Using a purely visual presentation might have less effects on the results.

Coming to the technical limitations and abnormalities. There were issues regarding the resolution. Also, the environment in which the character was placed could have affected the results. Familiarity with VR was also an issue. Some participants stated it was their first time being in VR, others had lots of experience. This could have lead to possible effects on the perception of the virtual avatar which was uncanny to some of the participants.

6 CONCLUSION AND FUTURE WORK

In this paper we presented a study comparing the empathy towards a 2D video-call like conversation and a realistic 3D VR model. The VR condition was described by participants as real and authentic, whereas the 2D condition was described as distant. Further research is needed to determine if this is due to the novelty of the VR technology or whether presence and realness played a corresponding factor.

Further research on the effects of gender on the empathy for the character should be conducted. We recommend focusing on the VR conditions and splitting the condition into using a male and a female avatar with the same scenario and same facial expressions. This should answer the question whether VR has a bigger influence on the empathy of males. This research also indicated that less empathic people enhance more empathy in VR, which is a good base for further investigations.

REFERENCES

[1] Laurie Batchelder, Mark Brosnan, and Chris Ashwin. 2017. The development and validation of the empathy components questionnaire (ECQ). *PloS one* 12, 1 (2017), e0169185.

[2] Mark H Davis. 1983. The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of personality* 51, 2 (1983), 167–184.

[3] Mark H Davis. 1996. Empathy: A social psychological approach.

[4] Nyaz Didehbani, Tandra Allen, Michelle Kandalaft, Daniel Krawczyk, and Sandra Chapman. 2016. Virtual reality social cognition training for children with high functioning autism. *Computers in human behavior* 62 (2016), 703–711.

[5] Hanna Drimalla, Niels Landwehr, Ursula Hess, and Isabel Dziobek. 2019. From face to face: the contribution of facial mimicry to cognitive and emotional empathy. *Cognition and Emotion* (2019).

[6] Agneta Fischer and Ursula Hess. 2017. Mimicking emotions. *Current opinion in psychology* 17 (2017), 151–155.

[7] Tamara M Haegerich and Bette L Bottoms. 2000. Empathy and jurors' decisions in patricide trials involving child sexual assault allegations. *Law and human behavior* 24, 4 (2000), 421–448.

[8] Ursula Hess and Agneta Fischer. 2013. Emotional mimicry as social regulation. *Personality and social psychology review* 17, 2 (2013), 142–157.
[9] Charlene Jennett, Anna L Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International journal of human-computer studies* 66, 9 (2008), 641–661.
[10] Esperanza Johnson, Ramón Hervás, Carlos Gutiérrez López de la Franca, Tania Mondéjar, Sergio F Ochoa, and Jesús Favela. 2018. Assessing empathy and managing emotions through interactions with an affective avatar. *Health informatics journal* 24, 2 (2018), 182–193.
[11] Michelle R Kandalaft, Nyaz Didehbani, Daniel C Krawczyk, Tandra T Allen, and Sandra B Chapman. 2013. Virtual reality social cognition training for young adults with high-functioning autism. *Journal of autism and developmental disorders* 43, 1 (2013), 34–44.
[12] Dennis Krebs. 1975. Empathy and altruism. *Journal of Personality and Social psychology* 32, 6 (1975), 1134.
[13] Robert W Levenson and Anna M Ruef. 1992. Empathy: a physiological substrate. *Journal of personality and social psychology* 63, 2 (1992), 234.
[14] Alan K Louie, John H Coverdale, Richard Balon, Eugene V Beresin, Adam M Brenner, Anthony PS Guerrero, and Laura Weiss Roberts. 2018. Enhancing empathy: a role for virtual reality? , 747–752 pages.
[15] David L Neumann and H Rae Westbury. 2011. The psychophysiological measurement of empathy. *Psychology of empathy* (2011), 119–142.
[16] Stef G Nicovich, Gregory W Boller, and T Bettina Cornwell. 2005. Experienced presence within computer-mediated communications: Initial explorations on the effects of gender with respect to empathy and immersion. *Journal of Computer-Mediated Communication* 10, 2 (2005), JCMC1023.
[17] Daniel Roth, Carola Bloch, Josephine Schmitt, Lena Frischlich, Marc Erich Latoschik, and Gary Bente. 2019. Perceived Authenticity, Empathy, and Pro-social Intentions evoked through Avatar-mediated Self-disclosures. In *Proceedings of Mensch und Computer 2019*. 21–30.
[18] Nicola S Schutte and Emma J Stilinović. 2017. Facilitating empathy through virtual reality. *Motivation and emotion* 41, 6 (2017), 708–712.
[19] Jun'ichiro Seyama and Ruth S Nagayama. 2007. The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence* 16, 4 (2007), 337–351.
[20] Donghee Shin. 2018. Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in human behavior* 78 (2018), 64–73.
[21] R Nathan Spreng*, Margaret C McKinnon*, Raymond A Mar, and Brian Levine. 2009. The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment* 91, 1 (2009), 62–71.