

Prediction of Self Perception based on Dominance and Trustworthiness by using Reverse Correlation.

Students

Arianne Leijenaar – 6043941

Saskia van Driel - 4147677

Jeroen Kusters - 6033881

Supervisor

Loek Brinkman



ABSTRACT

Humans perceive the world by making impressions and inferences about themselves and others, often called social perception (SP). Additionally, the way people perceive themselves is called the self-image, which is a part of, and in interaction with, SP. Both SP and self-image have been investigated in previous studies, but the methods that were used felt short for several reasons. The relative new method called 'reverse correlation' (RC) compensates for these shortcomings and is therefore used and tested for these concepts in this study.

This study investigates to what extent placement of the self within a group, based on dominance and trustworthiness, can be predicted by the mental representation of the self-image, using RC. There are two major purposes in this study: (1) to investigate if RC is a valid method to visualize self-image and (2) to demonstrate if RC is adequate to predict self-perception.

In the first experiment, fifty-one participants made a mental representation of their self-image, called a classification image (CI), with the help of RC. Subsequently, they ranked themselves on a dominance and a trustworthiness scale. In the second experiment, thirty independent raters ranked the classification images on a dominance and trustworthiness scale. The rating scores of the participants and the independent raters were investigated for correlation. Subsequently, if a correlation was found, a prediction of the self-perception could be made.

However, this study found no correlation between the self-judgment scores of the participants and the ratings of the independent raters. This leads to the conclusion that with RC as it is used in this study, no valid self-image could be made based on dominance and trustworthiness. Also, the placement of oneself within a group based on dominance and trustworthiness could not be predicted with RC the way it is used in this study.

Keywords: Social perception, self-image, reverse correlation, prediction model

INTRODUCTION

Behaviour is affected by the way people perceive the world, themselves and others (Trémeau, Antonius, Todorov, Rebani, Ferrari, Lee & Javitt, 2016). The way humans perceive impressions or make inferences about others is called social perception (SP) (Aronson, Wilson & Akert, 2010). SP contributes to the social determination of individuals relative to others, within a given group. The self-image is in part determined by perception of others and by comparing the self to others in a group (Lewicki, 1983; Turner et al., 1994). In this study, self-image is seen as the way the self is perceived, including subjective perceptions about one's own physical and psychological state. The self-image is intrinsically important in and for social interactions and could not be seen apart from SP (Hornsey, 2008). Most research performed in this subject focuses on how SP, and thus self-image, is affected by different factors (Ward, 2012). It has not yet been researched if SP can be predicted through comparing self-image to others. This is an interesting contribution to the current knowledge. It will give deeper insight in and how people perceive social situations. This

would create possibilities in predicting people's behaviour, since perception in part leads to behaviour. Which in part could create opportunities for economic, political and other social applications.

This study provides a contribution to previous studies by creating this method that could predict perception of an individual in a social setting. This prediction is made on the basis of a relatively new method, called reverse correlation (RC), which is used for visualizing mental images (Ratner, Dotsch, Wigboldus, van Knippenberg & Amodio, 2014). This new method is used because methods in the past fell short for predicting SP. First of all, this is because these methods could only investigate the conscious experience of SP. However, SP is largely an automatic mechanism, taking no more than 100-ms to infer character traits from a face (Willis & Todorov, 2006, 594). Because it takes 500-ms to incorporate sensory information from the outside world into conscious experience, SP is in part an unconscious process (Bergström, 2016). It is unknown up to which extent unconscious processes are being processed by the conscious mind (Greenwald & Banaji, 1995). For that reason, previous methods are unfit to capture the totality of SP. Secondly, there is no way in determining what variables are important for SP, since researchers may not be aware of these variables (Todorov, Olivola, Dotsch & Mende-Siedlecki, 2015). To summarise, current methods cannot measure SP sufficiently because, besides the fact that unconscious processes are difficult to study, there are too many variables to take into account and this makes prediction of SP hard.

RC is a method that could be used to compensate for these shortcomings (Ahumada, 1996; Ahumada, 2002; Ahumada & Lovell, 1971). RC enables to visualize mental representations: cognitive images of objects or phenomena that are not present to the senses (McKellar, 1957). Therefore, mental images do not have to be reconstructed through description alone. Instead, RC approaches an actual image resembling the mental representation, called the classification image (CI). RC research has mainly focused on visual perception (Ahumada, 2002; Solomon, 2002), neurophysiology (Victor, 2005; Ringach & Shaply, 2004) and SP (Dotsch & Todorov, 2012; Dotsch, Wigboldus, Langer & van Knippenberg, 2008; Oosterhof & Todorov, 2008). During our study for this paper, an article that investigates the prediction of SP, using RC came out (Zhan, Garrod, Rijsbergen & Schyns, 2017). Since SP is successfully researched with RC and SP and the self-image are dependent on each other, RC could be a useful method to visualize the self-image. After visualizing the self-image by making a CI of the self-image, SP could possibly be predicted from this CI. In this study, SP is investigated by the placement of oneself within a group by comparing the self to others by using RC.

Furthermore, since the self-image is based on physical and psychological traits, it might be too broad to investigate the self-image in total. For this reason, this study is focused on two character traits. Important

factors in social interaction are dominance - “the degree of perceived ability to inflict harm” - (Fruhen, Watkins & Jones, 2015) and trustworthiness - “the degree of perceived intent to inflict harm” - (Fruhen, Watkins & Jones, 2015). These are the first characteristics within social interaction which people rapidly and automatically ascribe to one’s appearance. (Wojcizke, 2005; Dotsch & Todorov 2012). For that reason, this study focuses on these two traits of the self-image.

Our study investigates to what extent placement of the self within a group based on dominance and trustworthiness can be predicted by a CI of the mental representation of the self-image. To do so, the perception of the self is incorporated in a prediction model that infers the connection between the selfimage and dominance and trustworthiness ratings of others. RC has not yet been used as a method to investigate the self-image. For that reason RC should be validated as a proper method to visualize the self-image at first. The procedure used for this study is adequate for both investigating validation of RC as a method to visualize the self-image, and is in addition adequate to predict behaviour based on correlations between CIs and a perception task.

METHODOLOGY

Participants and design

The total study consisted of two experiments. We chose to work with one particular sex, because according to Sutherland et al. (2014), males and females have altered facial perception. Male faces are chosen instead of female faces for the practical reason that the stimuli used in earlier studies are male (Dotsch & Todorov, 2012) and because of this a larger bundle of evidence could be compared to this study when using male participants. In addition, women perceive faces differently than men (Lewin & Herlitz, 2002; Bruce & Young, 1986).

For Experiment 1, fifty-five healthy male participants were recruited. Four participants were excluded after a sanity check due to double or missing data, resulting in fifty-one remaining participants (Age M: 26.1, SD: 5.1). Experiment 1 consisted of a RC-task, which is a two images forced choice task to create a CI. The second part of the first experiment was a self-judged perception task for measuring the selfperceived dominance and trustworthiness. Finally the participants conducted the Rosenberg’s Self-Esteem Scale questionnaire (Franck, de Readt, Barbez & Rosseel, 2008) and the CES-D, a questionnaire of Noordelijk Centrum voor Gezondheidsvraagstukken (NCG) regarding depressive feelings (Bouma, Ranchor, Sanderman, & Van Sonderen, 1995).

For Experiment 2, thirty healthy male independent raters participated. One independent rater was excluded because this person defined his/her sex as ‘other’, instead of ‘male, resulting in 29 independent raters (Age

M: 26.9, SD: 9.5). This experiment consisted of a rating task of the CIs generated in Experiment 1 and a short questionnaire for gathering personal information of the independent raters.

Stimuli

For the first task of Experiment 1, the two images forced choice RC-task, a base face (figure 1, image A) and randomly generated noise (figure 1, image B) from the study from Dotsch & Todorov (2011) was used. This variant of RC uses 4094 parameters in order to infer a random noised face pattern. For every randomly generated noise, an inversed noise was generated as well. Combining the noise, or its inverted noise, with the base face, resulted in a distorted unique variant of the face. By putting the unique newly created faces next to each other, a two images forced choice task could be conducted (figure 1, image C).

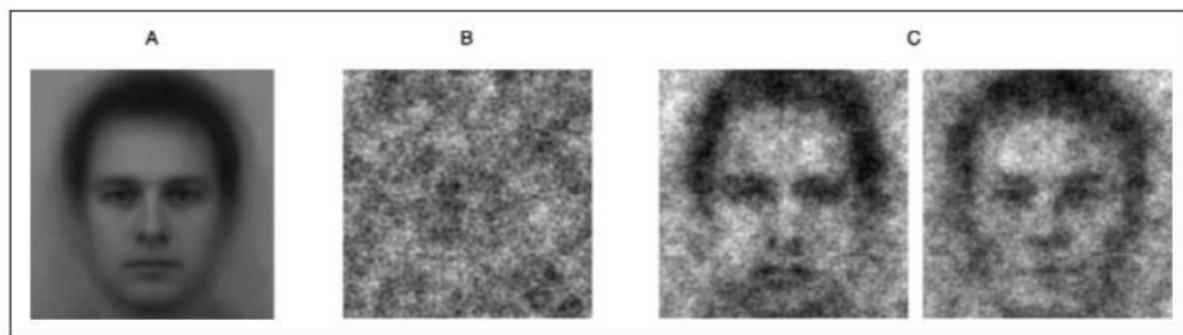


Figure 1. Base face 1 (A), random noise example (B), and example stimuli of noise superimposed on a single base face (C). The left stimulus shows that the base image with original noise superimposed and the right stimulus shows the base image with the negative superimposed. Adapted from “Reverse Correlating Social Face Perception” by R. Dotsch & A. Todorov, 2012, *Social Psychological and Personality Science*, 3, p. 563. Reprinted with permission

Every choice of the RC -ask was reported. With help of the program ‘R-Studio’ (version 1.0.143) and the ‘RCICR package’ (Ron Dotsch, 2016. RCICR: Reverse correlation image classification toolbox) the three hundred chosen noises were averaged and with combination of the base face, a classification image (CI) was created. By creating the mathematical opposite of the averaged noise, the inverse of the averaged noises were created. When these were combined with the base face, the anti-CIs were created. Those CIs and anti-CIs were used for Experiment 2.

Moreover, for the self-judged perception task two scales were created by using seven faces from the faceGen database (consulted in April 2017). Those pictures varied in equal steps from low to high dominance for the first scale and from low to high trustworthiness for the other scale.

Procedure

The data for this study was generated between 3th of May 2017 and 3th of June 2017. In the first phase of the analysis, the CIs were created. Both the participants and the independent raters performed the tasks

online. The experiments were created with the programming website gorilla.sc (www.gorilla.sc/about). Experiment 1 started with an introduction of the whole study and an informed consent. The experiment consists of two tasks. The first task, the RC-task, started with an introduction that contained a description of the procedure. Repeatedly a screen with two images was shown to the participants from which they had to choose as quickly as possible which image that they thought resembled them most by pressing “A” to select the left image and “L” to select the right image on their keyboard (figure 2, image A). The task consisted of 300 trials, separated by a fixation cross in the centre of the screen to protect from a bias towards the last chosen image because the attention could still be on the last image.

Hereafter, the participants had to perform the second task, the self-judged perception task. In this part of the experiment, participants needed to rank themselves two times (once for dominance and once for trustworthiness) on a scale of seven pictures ranging from not dominant/trustworthy to very dominant/trustworthy using a slider with a score from zero to fourteen (figure 2, image B).

Finally, the participants had to fill in a questionnaire regarding gender, age, marital status, living situation, education and employment, in order to possibly control for biases in a post hoc study, followed by twenty questions about mood conducted by the Dutch Centre of Health issues and ten questions about self-esteem based on the Rosenberg’s Self-Esteem scale (Franck, de Readt, Barbez & Rosseel, 2008) that is used for a study of Mitzy Kennis.

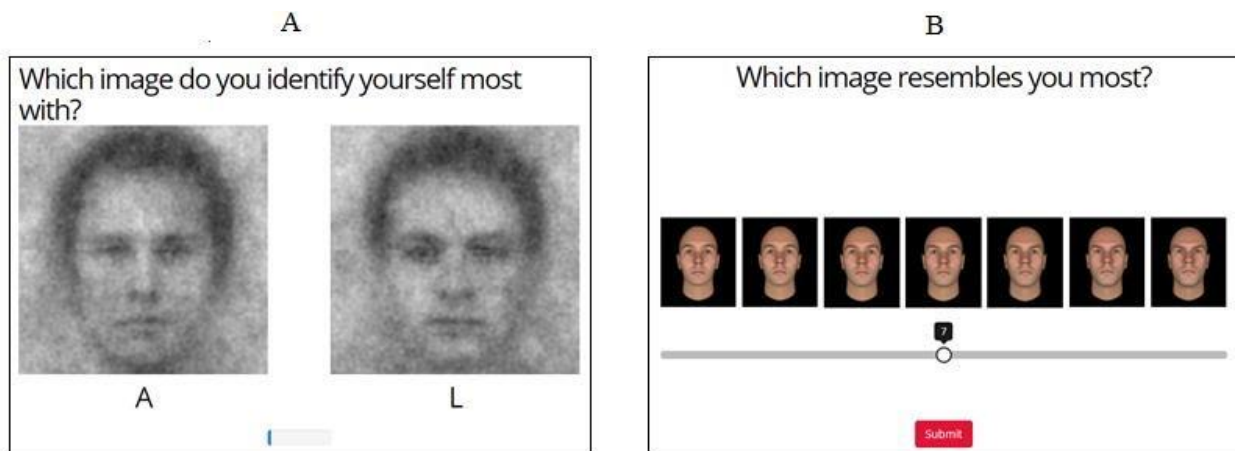


Figure 2: Example of task 1 of Experiment 1, the RC-task (A), and example of task 2 of Experiment 2, the self-judgement task (B)

For the second experiment, after agreement to the informed consent, thirty independent raters performed a scaling task. Because of the large number of trails, we decided to split the CI-dataset into two tasks. For that reason, fourteen independent raters rated the first 26 CIs and anti-CIs, fifteen independent raters rated the second 25 CIs and anti-CIs. They got instructions to judge how dominant or trustworthy a given CI or

anti-CI is by using a slider on the same scale of seven pictures ranging from not dominant/trustworthy to very dominant/trustworthy, described above. Again, the score ranged from zero to fourteen and the independent raters did not know the scales were based on dominance and trustworthiness. All the participants' personal CI and anti-CI were repeatedly shown one by one in random order (figure 3). An overview of the whole procedure is shown in figure 4.



Figure 3: Example of Experiment 2, the rating of the CIs and anti-CIs

Statistical Analysis

In order to validate RC as a method for presenting the self-image of dominance and trustworthiness, the rankings of the self-judged perception task are compared to the independent ratings of the CIs. If the correlation between the two is consistent and significant, RC is judged to be a valid method for representing the self-image. The significance level will be set on $\alpha = .5$.

Both the self-judged perception task scores and the rating score of the independent raters were imported into the statistical program 'SPSS' (version 22). The scores were calculated first calculated linearly, using a linear regression technique. Secondly, if this relationship was not found, the possible relationship was also checked on quadratic regression.

If a correlation results significantly to the regression coefficient 1, a perfect linear relation would be found. This would result in validation of the technique and in the possibility of predicting the self-perception. If this technique is not validated, because of a deviating value than 1, the self-perception could still be predicted if a significant correlation is found.

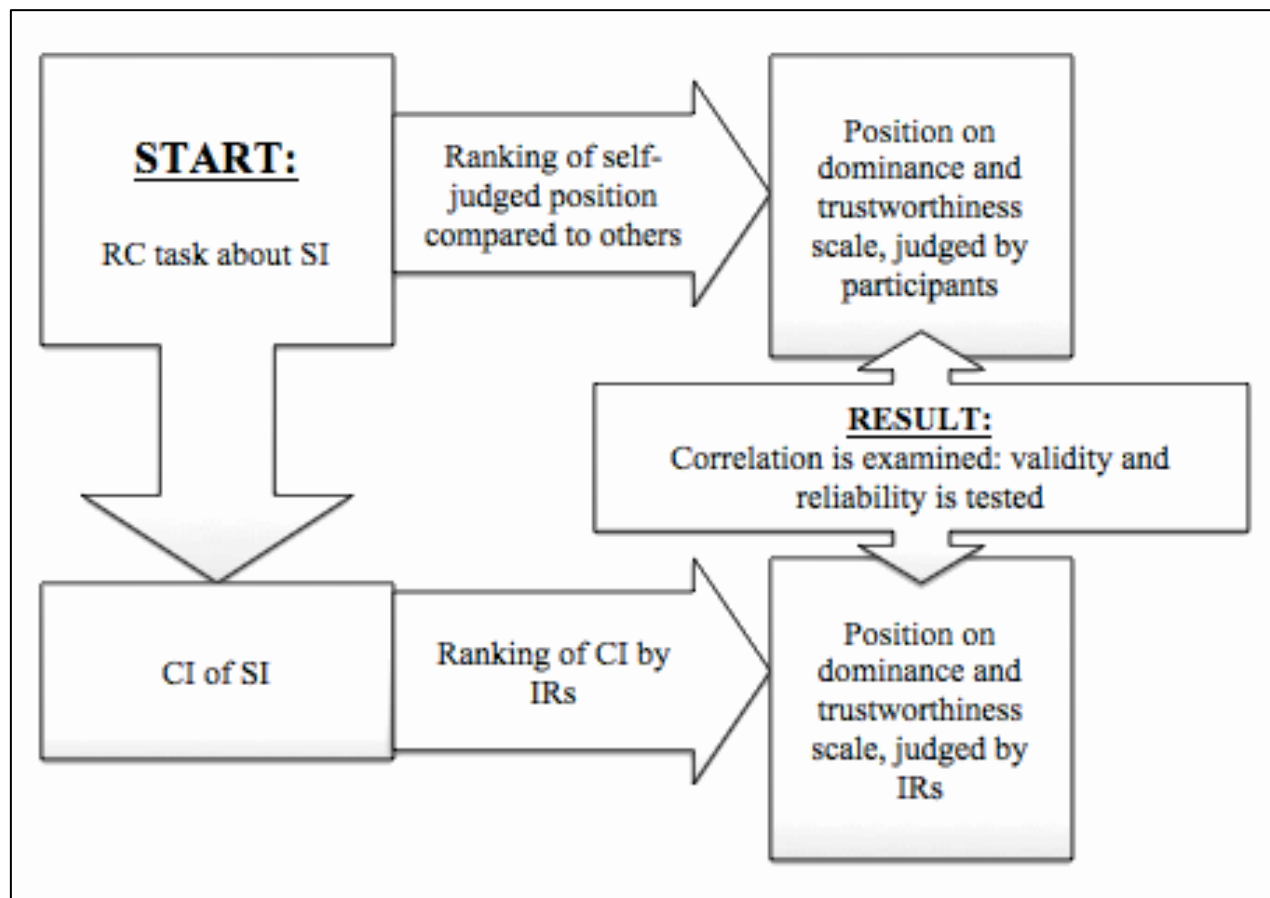


Figure 4: Overview of methodology

RESULTS

To test whether RC is a valid method to predict self-perception, the median of the dominance and trustworthiness scores from the 29 independent raters for each participant were conducted. These scores were compared to the dominance and trustworthiness scores of the participants self-judged perception task. An overview of the determinants is shown in table 1 below.

Table 1:

Overview data Experiment 1 and 2

Participants n = 51	M	SD
Age (years)	26.1	5.1
Self-rating Dominance	8.55	2.96
Self-rating Trustworthiness	6.06	2.35

Independent raters n = 29

Age (years)	26.9	9.5
Median Rating independent raters Dominance	6.99	1.35
Median Rating independent raters Trustworthiness	6.03	1.71

Note: An overview of the determinants of the participants and independent raters. n = Sample Size, M = Mean, SD = Standard Deviation.

The correlations between the self-rating scores (dependent) and the medians of the independent raterscores (independent) were calculated. Based on the results of the scatterplots (figure 5) and the linear regression, we concluded that indeed there is no significant relationship between the ratings, both for dominance and trustworthiness. The regression coefficient for dominance was $\beta_2 = .185$ with the significance score of $p = .557$. For trustworthiness the regression coefficient was $\beta_2 = .061$ with a significance score of $p = .758$. Both p-values are reaching above the cut-off score of $\alpha = .05$ and were for that reason regarded as non-significant.

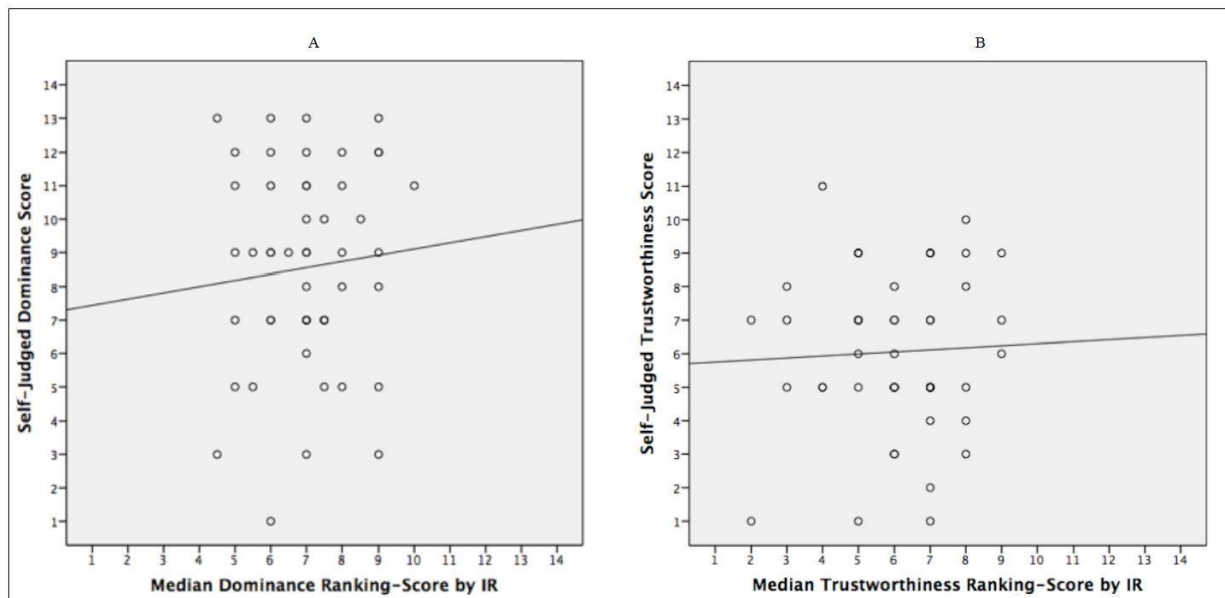


Figure 5: Scatterplot with the relationship between self-judged dominance and the median of the dominance ranking-score by independent raters (A), and scatterplot with the relationship between self-judged trustworthiness and the median of the trustworthiness ranking-score by independent raters (B)

In order to check for reliability of the independent raters scores, the inter-rater reliability of the independent raters was calculated. To do this, an intraclass correlation coefficient (ICC) was calculated, with a two-way mixed model and based on absolute agreement. In most research, a Cronbach's alpha below .70 is

considered as unreliable. Unfortunately, the Cronbach's alpha was below .70 for the dominance rankings (.61 for the first fourteen raters and .60 for the last fifteen raters) and for that reason is considered to be unreliable. The ratings on trustworthiness showed Cronbach's alpha of .82 for the first fourteen independent raters and .71 for the other fifteen independent raters and are counted as reliable.

Post hoc analysis

After post hoc analysis, we corrected the data of all the participants for lack of motivation. To test if a clearer relation could be found when unmotivated participants and raters are excluded, participants with a reaction time below five hundred milliseconds were excluded. This resulted in the exclusion of six participants.

To correct for unreliable raters, the raters with the biggest negative impact on reliability were deleted from the dataset until the inter-raters reliability score of at least .70 was reached. Six raters for dominance were excluded. This led to increase of the Cronbach's alpha from .61 to .70 for the first fifteen raters for dominance and from .60 to .71 for the last fifteen raters. New linear regressions were tested. For dominance, this led to the regression coefficient of $\beta_2 = .211$ with the significance score of $p = .495$. For trustworthiness the regression coefficient became $\beta_2 = .152$ with a significance score of $p = .481$. Both p-values are still higher than the cut-off score of $\alpha = .05$ and are for that reason insignificant. An overview of the new regressions can be found in table 2 and figure 6.

Table 2

Comparison ad hoc and post hoc analysis

Ad Hoc	β_2	p
Regression for dominance	.185	.557
Regression for trustworthiness	.061	.758
Post Hoc		
Regression for dominance	.211	.495
Regression for trustworthiness	.152	.481

Note: β_2 = regression coefficient, p = significance

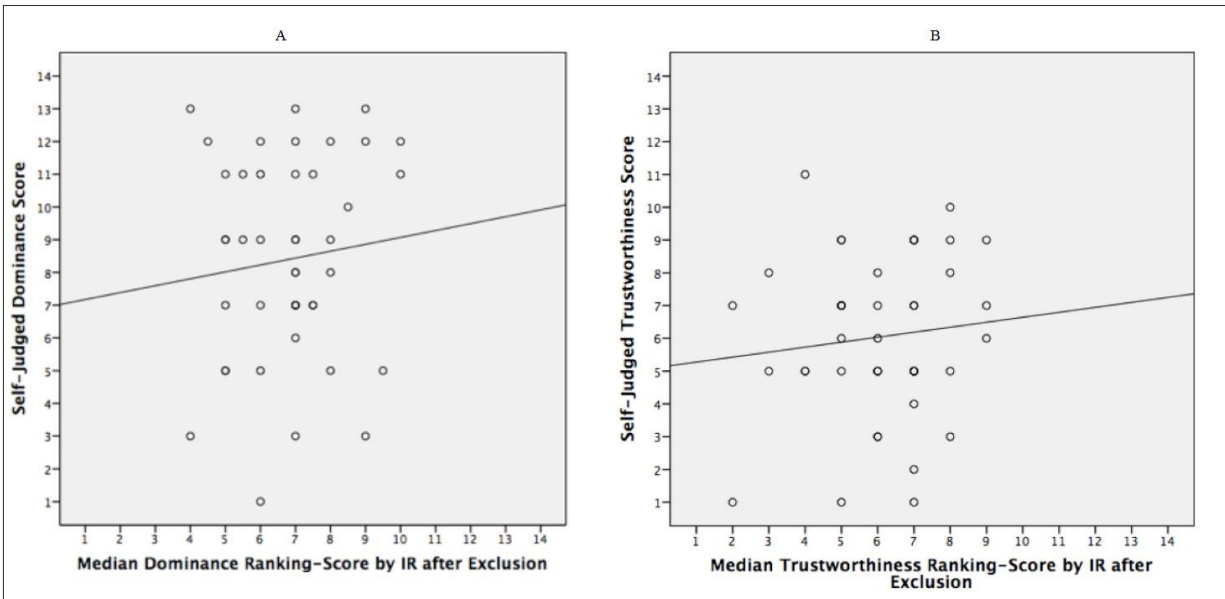


Figure 6: : Scatterplot with the relationship between self-judged dominance and the median of the dominance ranking-score by independent raters after exclusion of unreliable participants and independent raters (A), and scatterplot with the relationship between self-judged trustworthiness and the median of the trustworthiness ranking-score by independent raters, after exclusion of unreliable participants and independent raters (B).

A next possible reason why no linear relationship could be found is because the relationship could be quadratic. Both possible relationships are tested, for both dominance and trustworthiness, but no significant relationship was found. The p-value was even bigger, as can be seen in table 3.

Table 3

Comparison linear with quadratic regression

	β_2	p
Linear regression for dominance	.185	.557
Linear regression for trustworthiness	.061	.758
Quadratic regression for dominance	.001	.997
Quadratic regression for trustworthiness	-.017	.890

Note: β_2 = regression coefficient, p = significance

DISCUSSION

General discussion of results

In this study, we investigated to what extent placement of the self within a group, based on dominance and trustworthiness, could be predicted by the mental representation of the self, using RC. No significant relationship is found between the self-ratings and the ratings of the independent raters on dominance and trustworthiness scales. Therefore, it may be concluded that the way the method used in this study is not a valid way of visualizing the self-image based on dominance and trustworthiness. Besides that, we may conclude that placement of the self within a given group, with other words self-perception, regarding dominance and trustworthiness, cannot be predicted with the method used in this study.

There are several possible explanations why this relationship is not found. First because one may suggest that RC is not a correct method for correlating dominance and trustworthiness rates of both self-image and the interpreted image of that person. Nonetheless, both dominance and trustworthiness scales have been measured in the past with RC, which produced significant results. However, the dominance and trustworthiness scales used in this study could be inadequate. The faces might have not measured what they had to measure because they were for example too vague and not distinctive enough. Besides, the faces in the rating scales did not look like the CIs at all, which made comparison difficult.

Secondly, the non-significant correlation is most likely to be caused either by the procedure of the experiment or the statistical analysis. Since the latter is done with much care and consideration and with the consultation of experts in the field (Loek Brinkman, Ron Dotsch, Jelmer Zondergeld, Mitzy Kennis & others), it is suggested that the procedure of the experiment led to the unexpected non-significant results. In particular, it is more plausible that the non-significant results are due to the way the experiment is executed by the participants. For example, it may be that the participants and the independent raters did not complete the experiment with full concentration or honesty. This is likely, because at least a dozen participants and raters indicated that the experiments were too long and/or too boring or vague. People complained about not seeing differences between the shown pictures, what made choosing hard for them. This may have resulted in too random answers, because the participants just pressed 'A' or 'L' without any reason why.

In addition, the performance pressure for conducting the experiments with care may have been lower as opposed to other experiments, since they were filled in online at home instead of in a laboratory. Moreover, the participants did the experiment without receiving a compensation, which may have lowered their motivation to fill in the experiments as well. A clear indication for the low performance motivation is found in the data, which showed that more than two third (151/209) of the participants terminated the experiment

halfway during the RC experiments. Only 54 participants completed the total RC experiment. The rate of not completing the total experiment was the same for the scaling experiments (44/60).

To compensate for lack of motivation, we corrected post-hoc for random answers. This has been executed by excluding participants and independent raters using reaction times during the statistical analysis. People who responded repeatedly below five hundred milliseconds were excluded from the data. This resulted in a less non-significant correlation, but still far from significant. We considered five hundred milliseconds as a good margin to exclude unmotivated participants because you cannot look at and criticize two images good enough in this amount of time to make a reliable choice. After this correction, we also checked the reliability of the independent raters. There was only one independent rater excluded for the dominance task and this exclusion even led to a less significant correlation.

Suggestions for further research

Although our study did not produce a significant correlation between the ratings of dominance and trustworthiness of the self and the other, regarding dominance and trustworthiness, we have several suggestions to improve this study, This, in turn, may lead to significant results. If there would have been a significant correlation between the perception of the self and others regarding dominance and trustworthiness, using RC methodology, the perception of the self could have been predicted. In that case, one could infer that RC is indeed a correct method to measure the dominance and trustworthiness of the self. Once the dominance and trustworthiness of an individual can be inferred by a RC-image, the latter can be used for multiple conceptual and practical problems, for example for therapeutic settings, economical purposes, or political decision making. This is because the perception of the self is very important for the behavior of a person in a given social situation. In addition, within social neuroscience, a long-standing research questions about predictive coding theory could be answered using this technique and additional experimental manipulations. However, this is one of the multiple applications a significant result could have produced or can produce in the future.

There are several suggestions to improve this study. The first suggestion is to conduct the experiments in a fixed setting, for example a laboratory. In that way, participants might feel social pressure to fill in the experiments more honest and take their time for it. We chose not to do so due to time pressure of collecting a minimum of fifty participants in less than two weeks. It was easier for participants to perform the tasks online then to invite them to the laboratory in Utrecht. Also, our participants and independent raters groups were relatively small. Future research should try to include a larger sample of participants and independent

raters. The non-significant relationship that is found now, may be significant if the study includes more participants.

Other improvements to get more reliable results are to make the experiments less boring and less vague. This is also related to the lack of motivation. This problem may be overcome with clearer instructions of the method used. If participants know how the shown images are conducted, they understand where the differences in the faces come from, what may make the differences clearer for them, resulting in more convinced choices. Additionally, more explanation about the procedure may make the experiment less vague. The instruction “What image do you identify yourself most with”, can be interpreted in different ways. You can interpret the identification based on your outward appearance, but also based on your character. If participants knew that they had to fill the experiment in based on character, they may have filled it in differently. The same counts for the scaling tasks, participants did not know that they rated themselves for dominance and trustworthiness. It is possible that telling the participants this, maybe even by personal oral instruction instead of digital instructions, influences the outcomes of the study in a positive way. Moreover, paying the participants for their effort could increase their motivation to complete the whole experiment as well. Unfortunately, we did not have any rewarding materials at our disposal. Decreasing the number of trials, which is now set to three hundred, would not be an option. This would lead to an unclear or less significant CI. A less significant CI would mean a more neutral and thus harder to judge face.

Another way to get clearer and more significant relationship could be to change the dominance and trustworthiness scale. This could be done by changing the pictures of the scale in pictures that are more similar to the CIs, for example in grey and with a certain kind of superimposed noise. Another way to change the scale is by removing the pictures, and let the participants and raters score for dominance and trustworthiness intensity from zero to ten. A downside of this option could be that the participants and independent raters would already know on what basis the scales are constructed (i.e. dominance and trustworthiness). This could lead to bias, because people could give higher scores to themselves than to others. This is not a problem if this is done consistently for the predictions, but it would give a bias for the validation of RC for self-image.

This experiment is conducted only by male participants and independent raters. Since male and female judge facial characteristics differently, it could be very interesting to see if the results of this particular study differ for male and female participants and independent raters. In that way, you could investigate if this method could be used for female as well, and if they really judge in a different way than male do. Next to that, it can be investigated if women judge themselves differently. If this is the case, it may be that females

perceive the world differently, resulting in different behaviour (Lewin & Herlitz, 2002; Bruce & Young, 1886).

CONCLUSION

All in all, we can conclude that this explorative study could be changed in many ways, and may be the beginning of other studies on the same topics in many ways. RC in combination with predictions, SP and self-image opens up a whole new world of possible studies that could and should be investigated because there is a lot to discover to improve our knowledge about how perception influences behaviour.

REFERENCES

- Ahumada, A. J. (1996). Perceptual classification images from Vernier acuity masked noise. *Perception*, 1831-1840.
- Ahumada, A. J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision*, 121-131.
- Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, 1751-1756.
- Aronson, E., Wilson, T., & Akert, R. (2010). *Sozialpsychologie*. Hallbergmoos: Pearson.
- Beard, B. J., & Ahumada, A. J. (1998). A technique to extract relevant image features for visual tasks. *Proceedings of SPIE Human Vision and Electronic Imaging III* (pp. 79-85). San Jose: SPIE.
- Bergström, F. (2016). *The Neural Substrates of Non-Conscious Working Memory*. Umeå: the Dean of Medical Facultle.
- Bouma, J., Ranchor, A. V., Sanderman, R., & Van Sonderen, E. (1995). Het meten van symptomen van depressie met de CES-D. Een handleiding. *Groningen, Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen*, 2-24.
- Bruce, V. & Young, A., 1986. Understanding face recognition. *British Psychological Society*, 77(3), pp. 305-327.
- Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, 562-571.
- Dotsch, R., Wigboldus, D. H., Langner, O., & van Knippenberg, A. (2008). Ethnic outgroup faces are biased in the prejudiced mind. *Psychological Science*, 978-980.
- Franck, E., de Readt, R., Barbez, C. & Rosseel, Y. (2008). Psychometric properties of the Dutch Rosenberg Self-esteem Scale. *Psychologica Belgica*, 48, pp. 25-35.

Fruhen, L., Watkins, C. D. & Jones, B. C. (2015). Perceptions of facial attractiveness, dominance and trustworthiness predict managerial pay awarded in experimental tasks. *Leadership Quarterly*, 26(6), pp. 1005-1016.

Cauldron (2017). Gorilla (version 20170526) [Software] Retrieved from <http://Gorilla.sc>.

Greenwald, A. G. & Banaji, M. R. (1995). Implicit Social Cognition: Attitudes, Self-Esteem., and Stereotypes. *Psychological Review* 102(1), pp. 4-27.

Hornsey, M. J. (2008). Social Identity Theory and Self-Categorization Theory: A Historical Review. *Social and Personality Psychology Compass*, 2(1), pp. 204-222.

Lewicki, P. (1983). Self-Image Bias in Person Perception. *Journal of Personality and Social Psychology*, 45(2), pp. 384-393.

Lewin, C. & Herlitz, A., 2002. Sex differences in face recognition – Women’s faces make the difference. *Brain and Cognition*, 50, pp. 121-128.

McKellar, P. (1957). *Imagination and Thinking: A psychological Analysis*. New York: Basic Books.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis for face evaluation. *Proceedings of the National Academy of Sciences of the USA* (pp. 11087-11092). Princeton: Princeton University.

Ratner, K. G., Dotsch, R., Wigboldus, D. H., van Knippenberg, A., & Amodio, D. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes and behavior. *Journal of Personality and Social Psychology*, 897-911.

Ringach, D., & Shaply, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science*, 147-166.

Solomon, A. J. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision*, 105-120.

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M. & Young, A. W., 2012. Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, 127, pp.

105-118.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, 455-460.

Todorov, A., Olivola, C. Y., Dotsch, R. & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, 66, pp. 519-545.

Trémeau, F., Antonius, D., Todorov, A., Rebani, Y, Ferrari, K., Lee, S. H., Calderone, D., Nolan, K. A., Butler, P., Malaspina, D. & Javitt, D. C. (2016). What can the study of first impressions tell us about attitudinal ambivalence and paranoia in schizophrenia? *Psychiatry Research*, 238, pp. 86-92.

Turner, J. C., Oakes, P. J., Haslam, S. A. & McGarty, C. (1994). Self and Collective: Cognition and Social Context. *Personality and Social Psychology Bulletin*, 20(5), pp. 454-463.

Victor, D. J. (2005). Analyzing receptive fields, classification images and functional images: Challenges with opportunities for synergy. *Nature Neuroscience*, 1651-1656.

Ward, J. (2012). *The Student's Guide to Social Neuroscience*. Hove: Psychology Press.

Willis, J. & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, 17(7), pp. 592-598.

Wojciszke, B. (2005). Morality and competence in person and self-perception. *European Review of Social Psychology*, 16, 155–188.

Zhan, J., Garrod, O. B., van Rijsbergen, N. J., & Schyns, P. G. (2017). *Efficient Information Contents Flow Down from Memory to Predict the Identity of Faces*. *bioRxiv beta*. Advance Online Publication.