

Challenging Physiognomy: Questioning the Idea that Facial Characteristics are Indicative of Personality

Yi Lin Wang (yi_lingo920@sinamail.com)

Department of Psychology
University of Southern Queensland, Toowoomba QLD 4350 Australia

Jeff Patrick (patrick@usq.edu.au)

Department of Psychology
University of Southern Queensland, Toowoomba QLD 4350 Australia

Abstract

Physiognomy; the idea that facial characteristics are indicative of personality has persisted within the science of psychology despite some questionable supporting evidence. Indeed the idea is not unreasonable if certain premise can be supported. The aim of this research was to test three related premise in order to ascertain whether people could accurately judge the personality of a stranger from only a superficial exposure. An experiment was devised which exposed participants to one of eight video clips. The video clips were all of the same person but varied in duration, whether the eyes were visible, and whether the person was talking. One hundred and forty participants took part in the study. After watching one of the video clips each participant was asked to assess the personality of the person in the video using a standard personality questionnaire. The null results challenge the findings of previous research in support of physiognomy.

Introduction

The aim of this study is to investigate how much information is needed to begin to form accurate perceptions of other people's personality. Specifically, this study will examine how three factors can affect the perception of personality where controls are placed on the perceived attractiveness of a target person and the emotional intelligence of the participant. The three factors are whether or not someone talks, makes eye contact, and the duration of exposure, and the specific personality traits are Extraversion, Conscientiousness and Core Self Evaluation.

The notion that facial characteristics are indicative of the personality of a person is called physiognomy (Collins, 1999). This idea existed in ancient Greece. In the Greek theatre festivals, different masks were worn by actors to represent different characteristic facial expressions (Brandt, 1980). For example, a "rustic" could be represented by "a broad lip and a snub nose" (Brandt, 1980, p. 93). This basic idea has persisted to the current day. A recent survey conducted by Hassin and Trope (2000) indicates that 75% of Israelis believed that it was possible to judge personal personality traits by looking at the face.

According to both Brandt (1980) and Terry (1989), Aristotle wrote the first known treatise on physiognomy. Aristotle linked the physical appearance of animals to persons, and he proposed people would have similar characteristics to particular animals if their physical appearance was similar to those animals (Brandt, 1980; Terry, 1989). During the late 18th and early 19th centuries, Johann Kasper Lavater published four volumes called "Essays on Physiognomy", which gave rise to a renewed interest in physiognomy (Brandt, 1980; Collins, 1999). Lavater (as cited in Collins, 1999) developed various principles for reading the face which were based on his own observations. He believed that people could learn to systematically describe another's personality through their facial features, by applying his rules.

However, Lavater's work was criticised by a number of scholars because his rules were based primarily on his own subjective opinion (Brant, 1980; Collins, 1999; Hegel, 1807, 1971). While early evidence in support of physiognomy was weak, the very idea that personality could be reflected in facial features provided a fundamental explanation of one of the ways people begin to form first impressions about others (Allport, 1949; Secord, 1969; Hartley, 2001). Nevertheless, the longevity and common popularity of physiognomy does not constitute scientific evidence.

Scientific studies have shown that there is a predictable relationship between facial features and observer judgement of personality traits (McCabe as cited in Secord, Bevan, & Duck, 1953; Secord & Bevan, 1956; Secord & Muthard, 1955). Some research even extends the concept of physiognomy to an investigation of the relationship between expressive behaviours (both verbal and nonverbal) and personality traits (Allport & Cantril, 1934; Borkenau & Liebler, 1992, 1995; Mobbs, 1968; Scherer, 1978). Their results also indicate a relationship between expressive behaviours and third party ratings of personality traits. Yet the fact that there is a reliable relationship between appearance and expression, and observer judgement, does not necessarily mean that the observer judgements are accurate.

The issue of the validity of physiognomic judgments is more controversial. Taylor, Peplau, and Sears (1994) point out that a person's perception may not be valid because people often judge another's characteristics based on stereotypes. Other earlier studies concluded little or no validity in physiognomic inference (Alley, 1988a; Cohen as cited in Alley, 1988a). In contrast, more recent studies that take both physical features and expressive behaviours like voice, eye contact and attractiveness into account; suggest that there may be some validity in perceivers' judgments (Albright, Kenny, & Malloy, 1988; Borkenau & Liebler, 1992, 1993; Kenny, Horner, Kashy, & Chu, 1992). These later studies concluded that two of the big five personality traits, namely Extraversion and Conscientiousness, are more accurately judged by a third party than other traits.

In summary, the validity of physiognomic judgment is still uncertain. Recent evidences suggest the judgments made based on physical features and expressive behaviours perhaps may be valid, but evidence is not sufficient to be absolutely certain. Nevertheless, there are sufficient grounds to formulate some expectations. Firstly, perception of personality would be more accurate when the target spoke. Secondly, the accuracy of perceived personality will be higher when the target maintains eye contact. Finally, the longer the duration of exposure, the higher the level of accuracy of perceived personality will be.

Methodology

Participants

136 participants took part in the study (100 female and 36 male). The age of participants ranged from 17 to 66.

Materials

Participants completed the Goldberg Unipolar Inventory (GUI) sub-scales for Extraversion and Conscientiousness, the Core-Self Evaluation Scale, the emotional intelligence scale, and the emotional intelligence scale.

The GUI is developed based on the structure of a five factor representation of personality. This inventory contains 100 unipolar items and 20 items for each factor (Goldberg, 1992). Due to the fact that only two factors, Extraversion and Conscientiousness, are measured in this study, 40 unipolar items were selected from the GUI for Extraversion and Conscientiousness dimensions. The coefficient alpha reliability for the sub-scales of Extraversion and Conscientiousness ranged from .90 to .92 and from .88 to .94, respectively depended on the study being reported. Every item was rated on a 1 (extremely inaccurate) to 9 (extremely accurate) Likert scale.

The CSES contains 12 items rated on a 5 point Likert scale (1 representing strongly disagree and 5 representing strongly agree). Half of the items were

positively worded, and half were negatively worded (Judge et al., 2003). The test-retest reliability and internal consistency for CSES are .81 and .84, respectively.

Schutte et al. (1998) developed a 33-item measure of emotional intelligence (EI) using a 5-point Likert scale provided for each item, ranging from 1 (strongly disagree) to 5 (strongly agree). This scale has high test-retest reliability and internal consistency, .78 and .90 respectively (Schutte et al., 1998).

The attractiveness of the target to observers is measured by a single question, "how attractive is the person in the video?" rated on a 5-point Likert scale from 1 (not attractive at all) to 5 (very attractive).

Procedure

One video clip had the target maintaining eye contact with the camera, but not talking, for two minutes. The second had the target maintaining eye contact and talking for two minutes. The target was restricted to talking about the content of a weather forecast, which was considered irrelevant to her personality. The confederate (target) was given the GUI of Extraversion and Conscientiousness and the self-rating form of CSES as a point of comparison for participant ratings of her.

Eight video clips were constructed from these two video clips in order to account for each of 3 independent variables. Half the clips showed the target talking, and half showed her silent. Half were 2 minutes long, and half were 30 seconds long. And finally, half had the eyes covered by a black stripe across the image, and the other half had no black stripe. Specifically, the 8 clips show the target:

- a) Having eye contact and talking for 2 minutes;
- b) Having eye contact and talking for 30 seconds;
- c) Not having eye contact and talking for 30 seconds;
- d) Not having eye contact and talking for 2 minutes;
- e) Having eye contact and not talking for 2 minutes;
- f) Having eye contact and not talking for 30 seconds;
- g) Not having eye contact and not talking for 2 minutes; and
- h) Not having eye contact and not talking for 30 seconds.

One of the eight different video clips were randomly allocated to each participants (perceivers). After viewing the video clip, perceivers were asked to complete the questionnaires in relation to the perceived characteristics of the target.

Results

A 2 x 2 x 2 between-subjects multivariate analysis of covariance (MANCOVA) had been conducted to investigate any main effects and any interaction effects of independent variables (IVs) on dependent variables

(DVs) after the effects of covariates on DVs were controlled. There were three DVs: the accuracy of perceivers' judgment (the differences between target rating and perceiver ratings) on Extraversion, Conscientiousness, and Core Self-Evaluation. Adjustment was made for two covariates (CVs), the attractiveness of the target perceived by perceivers and the emotional intelligence of the perceivers. Independent variables (IVs) were voice (talking or not talking), eye contact (having eye contact or not), and duration of exposure (30 seconds or 2 minutes).

The result indicates that the two CVs had a significant impacts on the combined DVs. The attractiveness of the target perceived by observers significantly relates to the accuracy of their judgment, $F(3, 124) = 5.97, p < .05, 1-\beta = .95$. The emotional intelligence of the perceivers also provided significant effect on the combined DVs, $F(3, 124) = 4.21, p < .05, 1-\beta = .85$. The strength of the association between the attractiveness of the target perceived by observers and DVs was small, partial $\eta^2 = .13$. An even smaller association is found between the emotional intelligence of the perceivers and DVs, partial $\eta^2 = .09$. These confounding effects were controlled.

The results of the MANCOVA showed that no interaction effects are significant. Among the three IVs, only eye contact has a significant main effect on the composite of DVs after removal of the effects of the CVs, $F(3, 124) = 3.95, p < .05, 1-\beta = .82$. The proportion of the variance of the combined DVs that is attributable to this significant effect is tenuous, partial $\eta^2 = .09$. Eye contact does appear to affect the accuracy of judgment on Extraversion, $F(1, 126) = 6.03, p < .05, \eta^2 = .05$, but does not significantly affect the accuracy of judgement on Conscientiousness and Core Self-Evaluation. However, there is not sufficient power to support the reliability of the effect of eye contact on the accuracy of judgment of Extraversion ($1-\beta = .68$).

In addition, the raw scores for three DVs are converted to the z-scores in order to compare the level of the accuracy of perceivers' judgments among DVs, as the rating scale for the traits of Extraversion and Conscientiousness is different from the scale for Core Self-Evaluation. To achieve this comparison, the percentages of the number of perceivers who obtain z-scores within the range of positive and negative 0.5 SD from the mean, were investigated for each DV. Given that a standardized distribution has a mean of 0 and standard deviation of 1, if the standardized score of perceiver's judgment is closer to 0, the accuracy of their judgment will be higher. Here, the z-score of 0.5 was employed as a cut-off score for distinguishing whether perceiver's judgment has some accuracy or not. Therefore, the proportion of perceivers with z-score less than 0.5 in a given personality trait indicates the number of people who can judge that trait accurately to some extent.

The results indicate that the percentages of perceivers with z-scores less than 0.5 for the accuracy of judgment on Extraversion, Conscientiousness and Core Self-Evaluation are 33.80%, 37.50% and 45.60%, respectively. The percentage of normal population expected to fall within 0.5 SD by chance is 38.30%, and the accuracy of judgment on Core Self-Evaluation are the only ones to exceed this parameter.

In short, this finding suggests that, regardless of whether or not perceivers have received vocal cues of the target, and whether the duration of exposure is 30 seconds or 2 minutes, the accuracy of perceivers' judgment does not demonstrate significant differences. Nonetheless, there is a reliable difference between the accuracy of perceivers' judgment and whether or not the target maintains eye contact with the perceivers. The results further suggest that the accuracy of judgment on Extraversion makes a significant contribution to this difference, but there is not sufficient power to support that this contribution is reliable. In addition to this, there are more people who make accurate judgments on the trait of Core Self-Evaluation than on Extraversion and Conscientiousness.

Discussion

This study examined the level of accuracy of perceivers' judgment on the personality traits for eight different conditions (whether or not vocal cues were provided to the perceivers, whether target maintained eye contact with the perceivers, and whether the duration of exposure to perceivers is 30 seconds or 2 minutes).

The findings are inconsistent with previous research (Blackman & Funder, 1998; Borkenau & Liebler, 1992). The study conducted by Borkenau and Liebler show that perceivers tend to have higher accuracy on their judgment of another's personality when they view that person in sound-film condition than in still-image condition. In addition, Blackman and Funder's study point out the accuracy of perceivers' judgment tend to increase as the duration of exposure rise. The inconsistent results between the current study and previous research may be due to the fact that previous research employed the correlations between target ratings and perceiver ratings as a criterion for the accuracy of judgment. In contrast, this study used a stricter criterion, the differences between target rating and perceiver ratings for the accuracy of judgment. Employing the target-perceiver correlation as a criterion for the accuracy of judgment may lead to inappropriate findings. This is because it is possible to obtain a high target-perceiver correlation, even if there is a big difference between the target's rating and the perceiver's rating, as the target-perceiver correlation tends to be high as long as the rating pattern between the target and the perceiver is similar. Therefore, the results found in Blackman and Funder's study may not correctly reflect the real effect of vocal cues and

duration of exposure on the accuracy of perceivers' judgments.

An alternative possibility for gaining inconsistent results with previous research may be due to Type II error. The sample size of this study is sufficient to detect a medium to large effect size, or strength of association between each IV and each DV, but it is not sufficient to detect a small effect size. Hence, if there are reliable but small effects between voice and duration of exposure, and the accuracy of judgment, this study will not be able to reliably detect the effects due to insufficient sample size. In addition, there is still another possibility for having no reliable effect for duration of exposure on the accuracy of judgment in this study. That is, this effect may only exist when the duration is greater than five minutes, as Blackman and Fuder's (1998) have previously demonstrated. In contrast, the finding that the voice cues provided did not aid judgement of personality is not entirely unexpected. Only two out of six experiments conducted by Allport and Cantril (1934) had a higher percentage of perceiver/target convergence on the trait of extraversion than one would expect by chance.

As expected, this study finds that perceivers who could see the eyes of the target do tend to have more accurate judgment than those who have not. Nonetheless, the reliable effect of eye contact is only found on the accuracy of judgment of Extraversion and there is not sufficient power to be certain that this apparent effect is not in fact a Type I error. The power of the analysis in this case was only 68%.

In addition, this study found that the number of perceivers who judge the Extraversion and Conscientiousness of the target reasonably accurately, is less than the number expected by chance. This finding is also contrary to previous research. Previous research points out that people can judge a stranger's personality traits, especially Extraversion and Conscientiousness, reasonably accurately (Borkenau & Liebler, 1992; Norman & Goldberg, 1966; Watson, 1989). The inconsistent findings may also be due to the use of different criteria as the accuracy of judgment between previous research and present study, as aforementioned. It is also possible that facial cues are not valid for judging the Extraversion and Conscientiousness of another person, but rather body features and body movement may be to source of valid cues about the person. In previous research, the information about targets that was revealed to perceivers included whole physical features and expressive behaviours, such as voice and nonverbal behaviour, but the present study only offered the information related to face and voice to perceivers. Hence, there may be a need to extend the definition of physiognomy from facial features to physical features and expressive behaviours.

Of particular interest is the finding that more perceivers judged Core Self-Evaluation more accurately

than Extraversion and Conscientiousness. This implies that Core Self-Evaluation may be associated with more observable cues than Extraversion and Conscientiousness. Given that the current study is the first to explore the accuracy of perceiver judgement of Core Self-Evaluation, these findings would need to be independently replicated in order to be sure of the authenticity of this effect.

One issue still remains, how can largely null results such as those presented here inform our understanding of how physical and behavioural characteristics may transmit information about one's personality? In this regard challenging physiognomy, and even allowing for extended definitions of physiognomy that include some interpersonal behaviours, is not the same as challenging phrenology. Physiognomy has enjoyed at least some scientific support, and the concept is appealing when we can more accurately perceive personality from extended interpersonal contact. Nevertheless, it is likely that earlier research incorrectly interpreted relatively weak correlations between perception of personality and the person's actual personality even when there may have been a distinct gap between the two.

The current study can at least rule out some potential explanations about how interpersonal perception operates. Perception of complex characterological constructs like personality are not apparently manifest given only superficial contact. This knowledge should allow researchers to refocus their efforts on more worthwhile agents such as patterns of behaviour cues through extended contact, inter-cultural differences in perception, gender difference in both expression and perception, and the complexities of interpersonal interaction.

Conclusion

The present study found that voice and duration of exposure do not affect the accuracy of judgment significantly, but eye contact seems to have some effect on the accuracy of judgment on Extraversion. Due to insufficient power to support the effect of eye contact on the accuracy of judgment of Extraversion, strictly speaking all three IVs, namely voice, eye contact and duration of exposure, do not significantly affect the accuracy of perceivers' judgments on personality. That is, the information revealed in the face and voice may not be valid for making accurate judgment about personality. Accordingly, there may be a need to extend the definition of physiognomy from facial features to whole physical features and expressive behaviours.

Acknowledgments

We would like to acknowledge and thank the Community and Organisational Research and Evaluation (CORE) Unit within the Faculty of Science at the University of Southern Queensland for their support and assistance with this project. We would also

Wang, Y. L. & Patrick, J. (2006). Challenging physiognomy: Questioning the idea that facial characteristics are indicative of personality. In M. Katsikitis (Ed.), *Proceedings of the APS Annual Conference: Bridging the Tasman: Science, culture and practice*. Melbourne: Australian Psychological Society.

like to acknowledge the support of the Department of Psychology.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55*, 387-395.
- Alley, T. R. (1988a). Physiognomy and social perception. In T. R. Alley (Ed.), *Social and applied aspects of perceiving faces* (pp. 167-186). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Allport, G. W. (1949). *Personality: A psychological interpretation*. London: Constable & Company.
- Allport, G. W., & Cantril, H. (1934). Judging personality from voice. *Journal of Social Psychology, 5*, 37-55.
- Blackman, M. C., & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgment. *Journal of Experimental Social Psychology, 34*, 164-181.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*, 645-657.
- Borkenau, P., & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology, 65*, 546-553.
- Borkenau, P., & Liebler, A. (1995). Observable attributes as manifestations and cues of personality and intelligence. *Journal of Personality, 63*, 1-25.
- Brandt, A. (1980). Face reading: The persistence of physiognomy. *Psychology Today, 90*, 90-96.
- Collins, A. F. (1999). The enduring appeal of physiognomy: Physical appearance as a sign of temperament, character, and intelligence. *History of Psychology, 2*(4), 251-276.
- Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55*, 149-158.
- Goldberg, L. R. (1992). The development of markers for the Big-Five Factor structure. *Psychological Assessment, 4*, 26-42.
- Hartley, L. (2001). *Physiognomy and the meaning of expression in nineteenth-century culture*. Cambridge: Cambridge University Press.
- Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 78*, 837-852.
- Hegel, G. W. F. (1971). *The phenomenology of mind* (2nd ed.) (J. B. Baillie, Trans.). London: George Allen & Unwin. (Original work published 1807)
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale: Development of a measure. *Personnel Psychology, 56*, 303-331.
- Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62*, 88-97.
- Mobbs, N. A. (1968). Eye-contact in relation to social introversion/extraversion. *British Journal of Social and Clinical Psychology, 7*, 305-306.
- Moore, B. (Ed.). (1999). *The Australian Oxford dictionary*. South Melbourne, Victoria: Oxford University Press.
- Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology, 4*, 681-691.
- Saklofske, D. H., Austin, E. J., & Minski, P. S. (2003). Factor structure and validity of a trait emotional intelligence measure. *Personality and Individual Differences, 34*, 707-721.
- Scherer, K. R. (1978). Personality inference from voice quality: the loud voice of extroversion. *European Journal of Social Psychology, 8*, 467-487.
- Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J. et al. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*, 167-177.
- Secord, P. F. (1969). Facial features and inference processes in interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 300-315). Stanford, California: Stanford University Press.
- Secord, P. F., & Bevan, W. (1956). Personalities in faces: III. a cross-cultural comparison of impressions of physiognomy and personality in faces. *Journal of Social Psychology, 43*, 283-288.
- Secord, P. F., & Muthard, J. E. (1955). Personalities in faces: IV. a descriptive analysis of the perception of women's faces and the identification of some physiognomic determinants. *Journal of Psychology, 39*, 269-278.
- Secord, P. F., Bevan, W., Jr. & Dukes, W. F. (1953). Occupational and physiognomic stereotypes in the perception of photographs. *Journal of Social Psychology, 37*, 261-270.
- Taylor, S. E., Peplau, L. A., & Sears, D. O. (1994). *Social psychology* (8th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Terry, L. (1989). *About faces*. New York: Doubleday.
- Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57*, 120-128.