NO CONFEDERATES NEEDED: SOCIAL COMPARISON WITHOUT COLLABORATION

KAZUO MORI
Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

HIDEKO MORI
Bunka Women’s University Nagano Technical College, Nagano, Japan

A new experimental procedure for conducting social comparison experiments was developed to create artificially good and bad performers without the use of confederates. Anagram tasks of 2 different difficulty levels were presented using the fMORI technique (Mori, 2007) such that 2 tasks were viewed separately by 2 groups of viewers wearing different types of polarizing sunglasses. Those shown easier tasks would unwittingly perform those tasks better than the other group. Administration of the new procedure to replicate that used by Alicke, LoSchiavo, Zerbst, and Zhang (1997) with 40 mutually acquainted Japanese undergraduates showed that the new paradigm successfully created good and bad performers as expected. No participant noticed the trick. The results also showed that the participants attributed their performance to their own ability.

Keywords: social comparison experiments, confederates, fMORI technique, anagram tasks.
In the past, social psychologists have relied on confederates for experimental research involving social discrepancy among participants. For example, Asch (1956) used a group of confederates who responded incorrectly to a series of simple line judgment tasks, thus showing that a considerable proportion of the responses made by a naïve outnumbered participant conformed to the incorrect majority responses. There have been numerous similar examples in social psychological research, much of which has involved the use of confederates.

However, there are many drawbacks to using confederates. The most intrinsic problem of using confederates is that it inevitably gives a somewhat artificial feeling and causes participants uneasiness. Therefore, it is crucial to find confederates that can behave naturally and avoid being detected. These difficulties have prevented researchers from manipulating interpersonal relationships in their experiments effectively. It is relatively easy to set up a social comparison between participants who are unacquainted with each other; namely, a naïve participant and an unknown confederate. However, it is almost impossible to investigate social comparison between intimate friends or family members unless the experimenter secretly asks some of them to collaborate in advance. Research in the field of social psychology is more often conducted between friends or acquaintances than among strangers. Therefore, the manipulation of interpersonal relationships presents certain difficulties when conducting research in social psychology.

Recently Mori (2007) devised a new presentation method called the fMORI technique, in which two different visual images are simultaneously projected onto one screen using a single projector. This allows two groups of viewers to watch different images on the same screen at the same time, while each group is unaware of the presence of the other image. The fMORI technique allows the creation of good and bad performers in social comparison experiments by presenting either easier or more difficult tasks secretly and simultaneously to naïve participants. Participants who view easier tasks naturally outperform those who are shown more difficult ones.

In order for participants to believe that they have been given the same cognitive tasks, despite the fact that they are actually different, the tasks should have one common correct answer. It is also desirable that the answer be simple enough that even those performing the difficult task can easily find the correct answer if they look for it after having been told the correct answer. Anagram tasks are well suited to these requirements. If “PSITDU” and “STIPUD” are presented separately to two participants, the person presented the latter anagram can easily find the right answer, “STUPID,” while the person presented the former anagram should find it more difficult to get the correct answer; however, on hearing the right answer he/she should instantly recognize it. By utilizing this presentation trick and two-difficulty-level anagram tasks, Mori and Uchida (2009) caused lower-middle
level students to perform anagram tasks well above average and found that their self-efficacy was raised considerably after their contrived success.

People are involved in various social comparisons in everyday life. Some are undesirable ones that might damage their confidence or self-esteem. Alicke, LoSchiavo, Zerbst, and Zhang (1997) assumed that, when a person has been unambiguously outperformed and cannot dispute or distort the validity of the comparison, he/she is likely to exalt the outperformer by elevating the ability of the person whose performance dwarfs their own. In this way people can maintain their sense of competence while magnanimously acknowledging the superior attributes of the outperformer. Alicke et al. (1997) had pairs of participants (one of them a confederate) take a kind of perceptual intelligence test together. The participants were either to outperform the confederate or be outperformed and afterwards were asked to evaluate the intellectual ability of their counterparts (the confederates). As anticipated, those who were outperformed tended to exaggerate the ability of the outperformers in order to protect their own self-images against unfavorable social comparisons. The researchers dubbed this tendency to exaggerate an outperformer’s ability the “genius effect.” The genius effect was repeatedly observed in follow-up studies (Koehrsen, Abraham, Alicke, & Berger, 2003; Lassiter & Munhall, 2001).

When conducting this, our aim was to reproduce experimental manipulations in order to artificially create winners and losers, like those in the Alicke et al. (1997) study, without using confederates. Social comparison effects between two acquaintances are likely to be different than those between strangers. However, in typical confederate studies (such as in the Alicke et al. study and all of the follow-ups), the participants were not acquainted with the confederates. Therefore, only social comparison effects between unacquainted pairs could be examined. In the present study, by utilizing the new experimental procedure, we were able to conduct an experiment with groups of mutually acquainted participants rather than strangers.

**METHOD**

**PARTICIPANTS**

Ten mutually acquainted foursomes of Japanese undergraduates (five male foursomes and five female foursomes) participated in the experiments. The gender variable was not discussed in the present study as per Alicke et al. (1997).

**PROCEDURE**

*Random role assignment* Participants enlisted for what they believed was a preliminary study of the development of a novel psychological test. They were
invited to come to the psychology laboratory in groups of four of the same gender. On arrival, they were randomly assigned to one of the following roles: one performer of easy tasks (PE), one performer of difficult tasks (PD), and two observers (OE & OD).

**Anagram task instruction** The four participants were seated in two rows of two chairs placed side by side approximately one meter in front of the rear screen. The front seats were for the performers. The sunglasses suitable for each role were placed on each seat. The participants were told to wear the sunglasses to avoid glare. They were then told that they would be given several word puzzles presented one-by-one on the screen in front of them. The performers’ task was to figure out the correct word by rearranging the sequence of letters shown to them and announce the answer as quickly as possible. The observers sitting directly behind the performers were told that their task was to observe the activity of the two performers and make a written record. The observers were given a check sheet to record the winner for each task.

**Anagram task presentation** Each anagram task was presented for 20 seconds or until one of the performers answered correctly, whichever came first, and the experimenter then proceeded to the next item. There were 24 anagram tasks with two practice items preceding them. Eight relatively easy items, eight relatively difficult items, and eight medium-difficulty items were chosen according to a difficulty assessment which had been conducted previously. It was designed such that the performers who were shown the easier version would answer about half of the test items and that the performers presented the more difficult version might accidentally figure out the correct answers before their counterparts only once or twice. It took about 10 minutes to complete the anagram task session.

**Intelligence rating** Immediately after the anagram task session ended, participants were asked to rate their own intellectual levels on a questionnaire form. The performers rated the winners and losers, including themselves, while the observers rated the winners and losers as well as themselves. Ratings were made on a 10-point scale ranging from 0 (extremely low) to 9 (extremely high). To attenuate any hesitation to rate other people’s intelligence directly, the crucial rating was intermingled among a variety of filler questions concerning their attitudes, motivations, and other personality traits. It took about 15 minutes to complete the questionnaire.

**Postexperimental interviews** After the questionnaire session, the experimenter first asked whether the participants had noticed any abnormality in the presentation or not. She then explained the presentation technique and the research purposes in order to obtain the informed consent of the participants.
CREATION OF WINNERS AND LOSERS WITHOUT USING CONFEDERATES

Preparation of anagram tasks  The crucial point of this study was in the presentation of the anagram tasks. In fact, two different series of anagram tasks were presented on the same screen, one consisting of easier configurations (e.g., “STIPUD”) and the other of more difficult configurations (“DUSPIT”), both having the same correct answer (“STUPID”). These two different letter sequences were projected onto the same screen by means of two differently polarized light beams so that one could only be seen through one type of polarizing sunglasses while the other could only be seen through the other type (Mori, 2007). Performer PE and Observer OE wore the same type of polarizing sunglasses, which allowed them to view the easier anagram tasks, while Performer PD and Observer OD wore the other type and could therefore view only the more difficult anagrams.

Two sets of 24-item anagram lists were prepared from a previous study (Mori, 1993); one for the easy tasks and the other for the difficult tasks. The words used for the anagrams were chosen from three syllabic Japanese words written using six different letters. In the Japanese language, phonemes are basically composed of a consonant and a vowel. Although the Japanese orthography is a combination of kanji and kana, any Japanese word can be written using the Roman alphabet. The easy anagram tasks were made by rearranging only two letters while the difficult tasks were the result of much more complex rearrangements. Examples of easy and difficult tasks for the correct answer “NARUTO” are “NURATO” and “ORUNTA”, respectively. The difficulty level of each anagram task was assessed beforehand by having 18 undergraduate participants complete all of them. The averages of the difficulty levels of the difficult and easy tasks in terms of rates of correct answers were 0.53 and 0.11, respectively.

Drawing of two anagrams as an overlapped image to be seen separately  Two anagram tasks were drawn as an overlapped image using Adobe Photoshop. The easier versions of anagram tasks were drawn in order to be seen in black letters on a green background, while the difficult ones were drawn so as to be seen in black letters on a magenta background. The two images were integrated into a single PowerPoint slide. In order to mask the not-meant-to-be-seen images completely, random dot noises were added. For the detailed procedure for preparing the materials, see Mori (2007).

Projection of anagram tasks  The presentation of the anagram tasks was done by means of a PowerPoint slide show on an Apple iBook controlled by the experimenter. The slides were projected by an LCD projector manufactured by EPSON (LP-700). It had the LCD panels aligned in such a way that the one for the green component was placed perpendicular to those for the red and blue (magenta) components. The projection was done on a rear screen, a pane
of ground glass of 80cm x 80cm. The actual size of a PowerPoint slide on the
screen was about 50cm on the diagonal, about the same size as a 20-inch monitor.
Two types of polarizing sunglasses, a pair of each, were prepared with each
one suitable for watching images either in green or in magenta. The sunglasses
were the same as those used in previous studies (Kanematsu, Mori, & Mori,
1996/2003; Mori & Mori, 2008). A diagram of the overall experimental setting
is shown in Figure 1.

![Diagram of the experimental setting.](image)

**RESULTS AND DISCUSSION**

**EFFECTIVENESS OF THE EXPERIMENTAL MANIPULATIONS**

Postexperimental interviews revealed that no participant had noticed the
presentation discrepancy. The performers who were presented easier tasks
outperformed their counterparts in all 10 groups. The average number of correct
responses by the PE performers was 11.00 while responses by the PD performers
was 2.10, out of 24 anagram tasks. No PD performer outperformed the PE
performer in any group. Therefore, it was concluded that the new experimental
procedure successfully created winners and losers without using confederates.

**INTELLIGENCE RATINGS**

There was a pair of observers within each group in the experiment, one observing
the easier tasks and the other observing the more difficult tasks. Despite the fact
that they observed different tasks, their ratings of the intelligence levels for the
winners and losers were almost equivalent. A two-way analysis of variance (ANOVA) revealed that there were no significant differences between the two observers (\(F_{(1,18)} = 0.06, ns\)), nor in the interaction (\(F_{(2,36)} = 0.34, ns\)), whereas the main effect of the targets (selves/winners/losers) reached significance (\(F_{(2,36)} = 21.40, MSE = 2.57, LSD = 1.04, p < .05\)). Therefore, the observers’ ratings were combined and treated without differentiation in subsequent analyses.

Figure 2 shows the means for participants’ intelligence ratings. We carried out statistical analysis of the within-subject variable of ratings of the winner versus the loser, and the between-subjects variable of the rater (winner, loser, or observer).

![Figure 2](image)

**Figure 2.** Intelligence ratings of winners and losers by winners, losers, and observers. (The bars with arrows show the results of participants rating themselves.)

A two-way ANOVA revealed that there were significant main effects as well as interactions. The main effect of raters revealed a significant effect, \(F_{(2,37)} = 4.17, p < .05\), showing that the observers, winners, and losers rated differently. As expected, the within-subject variable yielded a highly significant effect, \(F_{(1,37)} = 39.50, p < .01\), showing that the winner’s intelligence was rated as higher than the loser’s. These results are basically the same as those obtained by Alicke et al. (1997), which means that the new experimental procedure successfully reproduced the findings of Alicke et al. (1997) not only in terms of experimental manipulations, but in basic findings as well.

**No Genius Effects**

However, we also found some small but crucial differences between this study and that of Alicke et al. (1997). An interaction between the rater (winner or loser)
and the target being rated (winner or loser) was significant, but in a different direction from the Alicke et al. finding, $F_{(2,37)} = 6.96, p < .01$. The results showed that the losers rated their own intelligence ($M = 3.6, SD = 2.25$) as lower than did the observers ($M = 5.4, SD = 1.85$), though the differences did not reach the significance level because of relatively large standard deviations ($F_{(2,37)} = 2.41, ns$), while they rated the winners’ intelligence ($M = 7.4, SD = 0.66$) the same as the observers did ($M = 7.6, SD = 0.97$). Meanwhile, the interaction of those winners’ ratings was also contrary to that reported by Alicke et al. (1997). The winners rated themselves much lower ($M = 5.4, SD = 1.20$) than did the observers and losers, $F_{(2,37)} = 17.48, p < .01$, and they rated the losers ($M = 4.8, SD = 1.89$) in the same way as the observer ($F < 1$). That is, the genius effect was not observed in the present experiment. Instead, the present results are best described as a “modesty effect”, in that both the winners and the losers rated themselves lower than the others did.

It should be noted that, besides the presence or absence of confederates, there were additional differences between the experimental procedures in the present study and those of Alicke et al. (1997). Closer examination of those experimental differences might provide an explanation of why no genius effects were observed in the present experiments. Among them, the most important difference may well have been the cross-cultural difference between the two sets of participants, Americans in Alicke et al. (1997) and Japanese in the present study. Numerous studies have been carried out concerning cultural differences in the social behaviors, attitudes, personalities, etc., of Americans and Japanese. See Kitayama and Markus (1994), for a review and summary of studies in this area.

Another important difference between the two studies was the interpersonal relationship between the participants. In Alicke et al. (1997), owing to the need for confederates, participants rated themselves and the strangers or confederates, while in the present experiment the participants were all members of a peer group who knew each other very well. It seems reasonable to suspect that people use totally different strategies to maintain their self-esteem among strangers and among close acquaintances or peers.

There were yet more differences between the Alicke et al. (1997) study and the present one. In the Alicke et al. experiment, observers watched from behind a one-way mirror so the performers would be unaware of being observed. However, the observers sat directly behind the performers in the present experiment. The tasks used in both studies were also different; Raven’s (1965) progressive matrices were used in the Alicke et al. study, whereas anagrams were used in the present one. The scoring procedure was also different: the participants (a participant and a confederate) were given 10 minutes to complete the test and then exchanged test sheets to score each other in the Alicke et al. study, whereas the participants’ performances were scored instantly item by item during the
presentation of anagram tasks in the present study. Nevertheless, none of these procedural differences seemed to be a pivotal cause for the differences observed in the dependent variables.

**CONCLUSION**

A new experimental paradigm utilizing a presentation trick and anagram tasks of different levels of difficulty was used to replicate a social comparison experiment such as that of Alicke et al. (1997), but without using confederates. The participants did not notice the presentation trick and were not even suspicious about it. It successfully created good and bad performers regardless of their actual ability. The results of intelligence ratings revealed that the participants attributed their good or bad performances to their own ability. The observing participants rated the contrived good and bad performers accordingly. Therefore, it can be concluded that the new experimental paradigm could be used successfully in a social comparison experiment without the need for collaboration. It is especially advantageous because the new paradigm allows researchers to study social comparisons among a variety of interpersonal relationships, such as acquainted pairs, as in the present demonstration, which has not previously been possible with pairs consisting of a confederate and a participant. This technique should also make it possible to conduct social comparison experiments with children of various ages without using child confederates.

**REFERENCES**


