Psychology Science, Volume 49, 2007 (3), p. 223-237

Self-estimates of attention performance

CHRISTOPH MENGELKAMP & REINHOLD S. JÄGER [1]

Abstract

In research on self-estimated IQ, gender differences are often found. The present study investigates whether these findings are true for self-estimation of attention, too. A sample of 100 female and 34 male students were asked to fill in the test of attention d2. After taking the test, the students estimated their results in comparison to their fellow students. The results show that the students underestimate their percent rank compared with the actual percent rank they achieved in the test, but estimate their rank order fairly accurately. Moreover, males estimate their performance distinctly higher than females do. This last result remains true even when the real test score is statistically controlled. The results are discussed with regard to research on positive illusions and gender stereotypes.

Key words: concentration, attention, gender, stereotypes, self-estimation

[1] Christoph Mengelkamp, Reinhold S. Jäger, Centre for Educational Research, University of Koblenz-Landau, Bürgerstr. 23, 76829 Landau, Germany; email: jaeger@zepf.uni-landau.de

Introduction

In the last ten years, several studies (cf. table 1) examining self-estimation of academic intelligence have been published, but these studies do not include measures of attention. Therefore, the present study investigates whether findings from research on self-estimated IQ that refer to the operationalization of intelligence without the inclusion of processing speed tasks are true for attention performance, too. However, a generally acknowledged definition of attention is lacking to date (cf. Schmidt-Atzert et al., 2006). The criterions of attention tests for cerebro-organically sound testees suggested by Westhoff and Hagemeister (2005, p. 37f) could pose an alternative:

- The testee must react to simple stimuli according to rules that are easy to remember.
- Actions must be carried out deliberately and as quickly as possible.
- Speed and quantity of errors are measured.

How is intelligence linked to concentration?

One approach is that concentration is a component of intelligence. In the Berlin structural model of intelligence (Jäger, 1982, 1984), processing speed is also defined as a facet of intelligence and is operationalized by three task groups containing figural, verbal and numerical material in the Berliner Intelligenzstrukturtest [Berlin intelligence structure test] BIS (Jäger, Süß, & Beauducel, 1997), the scale BIS-ZS (digit symbol test) is even suited as a marking test for the factor attention (Schmidt-Atzert, Bühner, & Enders, 2006).

Another scientific position contends that intelligence is a prerequisite of intelligence. Carroll (1993, p. 547ff) points out that attention is required for any cognitive performance and, consequently, for intelligence performances. This view is supported in a study by Schweizer, Moosbrugger & Goldhammer (2005) that predicts intelligence from concentration via structural equation modelling. At latent level, concentration correlated at .57 with intelligence. Nevertheless, the authors alleged that attention should be retained as an independent construct. Attention is not regarded as a component of intelligence, but as a source for the more complex construct of intelligence. Similarly to the study by Schweizer et al., the confirmatory analyses in a further study (Schmidt-Atzert et al., 2006) found a correlation of $r = .34$ between intelligence and attention factors. The result was replicated in a second sample with regrouped tests and a correlation of $r = .40$.

Both of the cited studies demonstrate that although intelligence and attention can be regarded as two different constructs, they correlate to a small to medium extent. Thus, regardless of whether concentration is seen as a component or a prerequisite of intelligence, it seems to be appropriate to investigate if findings from research on self-estimation of intelligence can be transferred to self-estimation of attention, as is the focus of this article.

Self-estimation of attention

In the dissertation by Scholz (2006), three studies are reviewed that correlate data from questionnaires about attention with attention performance. Scholz reports no or only slight correlations between the questionnaires and the test performance. Accordingly, her own

questionnaire "Concentration in everyday life" did not correlate significantly with different measures of attention, either. Thus, performance in a test of attention and the experience of attention in everyday life seem to be different aspects of the construct attention (Scholz, 2006, p. 52). Nevertheless, she found gender effects for her questionnaire, i.e. men had lower scores in the questionnaire indicating that they feel less disrupted and more concentrated in everyday life than women do. In contrast to the studies by Scholz, we do not construct a questionnaire about concentration in everyday life, but asked for an estimation of test performance as done in many studies about self-estimation in the field of intelligence. As there are no studies to the authors' knowledge that report upon the self-estimation of attention performance asking participants to estimate their test score, the findings from research on intelligence will be summarized briefly to begin with.

Self-estimation of intelligence

To what extent can one estimate one's own intelligence? In a meta-analysis by Mabe and West (1982), a correlation of r=.34 between intelligence and self-estimation of intelligence was reported, both unweighted and weighted by sample size. Because the meta-analysis was carried out in 1982 and therefore includes studies up to this date only, we searched for recent studies. These studies are listed in table 1 and also support the findings in the meta-analysis. The correlation of .34 between self-estimation and test score of intelligence found in the meta-analysis and more recent studies can be seen as a typical value of self-estimation of intelligence.

As can be seen in table 1, different tests of intelligence and different kinds of self-estimation were used, but the effects of only few context variables on the accuracy of self-estimation have been examined. The results are that participants give more accurate estimates when rating their intelligence by stating whether they can perform certain tasks on a percentile rank scale than when rating it on a scale labelled from "well" to "badly" (Holling & Preckel, 2005). Thus, social comparison seems to increase the accuracy of ratings, a finding that is also reported by Mabe and West (1982) in their meta-analysis. In contrast, neither using an anchor test before self-estimation nor giving feedback about the results of the anchor test before self-estimation altered accuracy in self-estimation (Holling & Preckel, 2005). In a study by Försterling and Morgenstern (2002, study 1), feedback was manipulated experimentally. The results show that unrealistically positive feedback leads to higher self-estimations of intelligence than unrealistically negative feedback. The self-estimations for realistic feedback and inverse, unrealistic feedback lay somewhere in between and did not differ from the other experimental conditions. Therefore, feedback seems to have an impact on self-estimation if it is unrealistic, but not if it is realistic. Beyond, efforts to increase correlations with the aid of indirect self-estimation, aggregation of items into scales, and weighting of items with their diagnosticity rated by experts did not lead to higher correlations between self-estimation and intelligence (Paulhus, Lysy, & Yik, 1998).

In the study by Rammstedt & Rammsayer (2002), self-estimates for Thunstone's seven primary abilities were obtained and correlated with test scores in these abilities. Significant correlations were found for verbal comprehension (r = .39), mathematical intelligence (r = .35), spatial intelligence (r = .27), reasoning (r = .22), and memory (r = .15). The corre-

Table 1:
Results of recent studies regarding the correlation between self-estimation of intelligence and performance measures of intelligence

Study	Sample	Intelligence test	Self-estimation	r
Holling & Preckel (2005) [a]	senior pupils, some university beginners	Intelligenz Struktur Test (IST-70) [Intelligence Structure Test]	percentage scale vs. rating scale from 1 to 9 with labels "well", "moderately", "badly"	.46
Försterling & Morgenstern (2002, study 1) [b]	students	4 scales from the Intelligenz Struktur Test (IST-70) [Intelligence Structure Test]	rating scale from 1 to 7 with pole labels "not gifted" and "very gifted"	.39
Furnham & Rawles (1999)	students	spatial ability and mental rotation (S&M Test), Part II	self-estimated IQ after explanation of the distribution characteristics of IQ	.16
Borkenau & Liebler (1993)	population, average age 34.0 years	Leistungsprüfsystem (LPS) [Performance Testing System}	rating scale from 1 to 7 with pole labels "not at all intelligent" and "highly intelligent"	.32
Robins & Beer (2001, study 2) [c]	students	Scholastic Assessment Test (SAT) & Grade Point Average (GPA)	8 items (α = .79), 3 of which refer to the comparison group of fellow students; 3 items inquire of GPA scores	.33 .22
Paulhus et al. (1998) [d]	students	Wonderlic Personnel Test	various scales for indirect and direct measurement of intelligence	.26 .24
Fingerman & Perlmutter (1994) [e]	population in 4 age groups from 20 to over 80 years	Watson Glaser Reasoning Tasks, Fuld Memory Test, Wechsler Adult Intelligence Scale (WAIS)	5-point Likert scale: "very poor, poor, good, very good, excellent"	.38
Rammstedt & Rammsayer (2002) [f]	students and vocational school pupils	Leistungsprüfsystem (LPS) [Performance Testing System], Berliner Intelligenzstrukturtest BIS [Berlin Intelligence Structure Test], Wilde Intelligenztest (WIT) [Wilde Intelligence Test]	7 visual analog scales for each of Thurnstone's primary mental abilities, labelled as: "extremely low, low, average, high, extremely high intelligence"	.18
Paulhus & Morgan (1997) [g]	psychology students	Wonderlic Personnel Test	3 items "creative, intelligent, wise" on a unipolar scale from 1 to 15 from "not at all" to "very much"	.20 .35

Note. [a] self-estimation scale experimentally varied; [b] feedback to students was experimentally manipulated in order to distort self-estimates; [c] first correlation at t_0, second correlation five years later; [d] direct (.26) and indirect items (.24) with no evident reference to intelligence employed for self-estimation; [e] mean of 36 reported correlations (9 tests* 4 age groups), lowest correlation -.29, highest .76; [f] mean of 28 reported correlations (7 tests * 2 education levels * 2 genders), lowest correlation -.19, highest .53; [g] first correlation after one week of group work, second correlation after seven weeks of group discussions (20 minutes once a week); uses scale according to personal message from D. Paulhus (October 6 2005)

lations for verbal fluency and perceptual speed were not significant. Therefore, not every trait can be estimated by the participants, and the question remains whether a correlation between self-estimated attention and actual attention in a test can be found at all.

Gender differences in the self-estimation of intelligence

When looking at self-estimation of intelligence, many studies found gender differences with men overestimating themselves more than women (Bennett, 1996; Furnham & Rawles, 1999; Hogan, 1978). Byrd and Stacey's (1993) study is an exception, as they found no gender difference regarding self-estimation, as is a study by Reilly and Mulhern (1995) in which the gender difference found was no longer significant after the exclusion of several outliers. In more recent studies, no difference was found regarding the estimation of intelligence in the sense of a common factor g; however, differences were found within facets of intelligence. Men's estimations of their logical (mathematical), spatial, and body-kinesthetic intelligence were higher than women's, but not those of their verbal, musical, interpersonal, and intrapersonal intelligence (Furnham, Fong, & Martin, 1999). The finding concerning mathematical intelligence and a factor that comprises mathematical and spatial intelligence was confirmed in another study (Furnham, Clark, & Bailey, 1999). A German study examining self-estimation of Thurstone's primary mental abilities and of four other domains of intelligence according to Gardner showed that men's estimations of their mathematical ability, spatial intelligence and reasoning were higher than women's, and vice versa, women's estimations of their musical and interpersonal intelligence were higher than men's (Rammstedt & Rammsayer, 2000). Similar findings were revealed in further studies in Germany. Men overestimated their reasoning ability and their spatial intelligence (Holling & Preckel, 2005; Rammstedt & Rammsayer, 2002). Moreover, Holling and Preckel discovered differences in memory, Rammstedt and Rammsayer in mathematical ability and speed of perception.

In the studies mentioned above, intelligence was estimated either without the actual implementation of an intelligence test, or before testing. When estimates were given retrospectively after the test, again, men overestimated their performance more than women did (Stankov, 1998). This was both true when testing the ability to find synonyms (inductive thinking) and also for Raven's Advanced Progressive Matrices. In another study, retrospective judgements of confidence for each item only revealed a tendency in the gender effect (Jonsson & Allwood, 2003). Accordingly, men's estimation of their ability to find synonyms tended to be more realistic than women's, who tended to underestimate themselves. However, both sexes overestimated their confidence concerning items about spatial intelligence equally.

One explanation for men's higher estimate of their intelligence could be that their performance regarding both fluid intelligence and general knowledge, a component of cristallized intelligence, is significantly better than women's, at least in some studies (Lynn, 1998; Lynn & Irwing, 2002). Above all, the finding concerning alleged superior male intelligence was supported in the area of spatial abilities (Kimura, 1999, cited from Lynn & Irwing, 2002; Masters & Sanders, 1993). Therefore, men's higher self-estimates could be based on actual, slight differences that are over-exaggerated during self-estimation (Furnham, Clark et al., 1999). If the actual intelligence score is partialled out of the self-estimates, however, the gender difference remains, excepting mathematical skills (Holling & Preckel, 2005;

Rammstedt & Rammsayer, 2002). Therefore, the finding that men estimate their intelligence as higher than women's might be accounted for in part by social stereotypes (cf. also Beloff, 1992). The fact that greater verbal abilities are ascribed to women and mathematical and spatial abilities to men does not change the stereotype whatsoever: "People who understand about calculus and geometry are clever. Anyone verbal is a chatterer" (Beloff, 1992, p. 309). The assumption that social stereotypes are the reason for differences in self-estimation is supported by the finding that men and women reach the same assessment when self-estimating interpersonal intelligence (cf. Gardner, 1998), where this stereotype does not apply (e. g. Bennett, 1996).

Effects of anonymity on self-estimation of intelligence

Beside the question of gender differences in self-estimation, another question in the present study concerns the context in which the self-estimate is made. In their meta-analysis, Mabe and West (1982) came to the conclusion that self-estimates with stated name are more accurate than anonymous self-estimates. The explanation given was based on Festinger's theory of cognitive dissonance (1954). They postulated that there are two motives that influence how accurate self-estimates are: (1) the desire for accurate self-estimation, and (2) the desire to protect or enhance one's competence. According to Mabe and West (1982), the former should apply in measurement situations in which an accurate self-estimate would not be disadvantageous to a person. This is particularly the case when the self-estimate remains anonymous. But reassuring the participant that his data will be handled confidentially should also conduce to the former need. According to the meta-analysis mentioned above, invalidating the anonymity of which participants are usually assured in scientific studies is enough to bring about differences in the accuracy of self-estimates.

Hypotheses

This study endeavours to replicate the cited findings on self-estimation of academic intelligence in the field of attention. Based on Mabe and West's meta-analysis and the reported newer studies on self-estimation of intelligence (cf. table 1), (1) we hypothesize that self-estimates of attention are possible, i.e. that self-estimation and actual performance correlate. If the findings from intelligence research can be transferred to attention, too, then (2) men should estimate their performance higher than women, even if the actual performance in the test is statistically controlled. Finally, the hypothesis from Mabe and West's meta-analysis will be examined, according to which (3) filling in the test anonymously will lead to more accurate self-estimates than under the non-anonymous condition.

Method

Sample

The sample comprises two subsamples of 79 students who were tested in the winter term 2005/2006 and 55 students who participated in the study in the summer term 2006. Both subsamples do not differ with regard to their attention score, $t(132) = .19$, $p = .851$, and their self-estimates are also equal, $t(132) = 1.10$, $p = .274$. Therefore, in the following, both subsamples will be combined in all following analyses. Thus, a total of $n = 134$ teacher students took part in the study. Of these, 100 (74.6 %) were female and 34 (25.4 %) were male. The mean age of the students is 22.8 years with a standard deviation of 3.9 years. The youngest student was 19 years old, the eldest 43.

Measures

To measure attention, the test of attention under pressure d2 was chosen because it is employed in various psychological disciplines such as clinical psychology, neuropsychology, traffic and transport psychology, educational psychology, and industrial and organizational psychology, and is one of the most frequently used methods in Germany. In the test d2, 14 lines must be processed. In each line, the letters d and p are printed with each up to two marks above or underneath the letters. All d's with two marks altogether must be crossed out within 20 seconds per line. After 20 seconds, the instructor tells the participants to proceed with the next line. The attention score "KL" is calculated from the number of d's crossed out correctly minus the number of letters crossed out incorrectly.

The students' self-estimation of their performance in the d2 was rated with regard to their performance compared with their fellow students: "Please estimate how well you did in the preceding test of attention. Mark how many percent of your fellow students came off worse than you." The answer scale ranged from 0 to 100 % in steps of five. Participants were not informed how their performance in the d2 test would be calculated. A retrospective comparison with fellow students was chosen because, according to the meta-analysis mentioned above, the highest correlations between performance in the test and self-estimate can be expected under these conditions (Mabe & West, 1982).

Procedure

The students were tested in one group at the beginning of the winter term 2005/2006 and of the summer term 2006 at the start of a lecture for teacher students. The participation in the study was voluntary. First, age and gender were ascertained, as were the first and surnames and also the e-mail addresses of half of the students. The other half of the students filled in the test anonymously and only supplied a self-generated code needed to receive feedback about the results. Then the test of attention under pressure d2 (Brickenkamp, 2002) was carried out with both groups together according to the instructions in the test manual. After completing the test, the students were asked to give an estimate of their test performance as described above. The students all received feedback about the number of processed tasks,

their amount of mistakes, their standardized attention score, their percentage compared with their fellow students, and the difference between their actual percentage and their self-estimate.

Results

All results were calculated on the basis of both the attention score "GZ-F" traditionally employed in the d2 test and the attention score "KL" recommended in the latest edition and described above. Both the attention score "GZ" and the attention score "KL" achieved a reliability of .95 (Cronbach's alpha)[2]. Because all results regarding both performance measures were identical apart from marginal deviations, only the results regarding the attention score "KL" recommended in the test manual will be reported in the following. In order to be able to discuss results opposing the hypotheses if necessary, all hypothesis testing will be carried out using two-tailed tests at a significance level of .05.

Initial analyses

First, the results of the attention score were compared to the norms for the age group from 19 up to 29 years given in the test manual: the mean attention score was $M = 183$ with $SD = 37$. In the present sample, the mean was $M = 196$ with $SD = 35$. Thus, the present sample performed one third of a standard deviation better than the norm. Further, the variance was not reduced compared with the population, and thus, there were no problems with restricted range.

Self-estimation of attention performance

As assumed in the first hypothesis, the self-estimates correlated significantly with the attention scores, $r(133) = .31$, $p = .000$. Thus, just under 10 % of the variance can be explained, a result that can be classified as a medium correlation (Cohen, 1988). Further, it is of interest whether the attention score is overestimated or underestimated. If the estimates were accurate, the mean estimated percentage should approximate to 50 %. However, the students underestimated their performance considerably at $M = 33.7$ % compared with the expected 50 %, $t(133) = -9.44$, $p = .000$, $d = .82$.

[2] According to the standards for educational and psychological testing (1999), test-retest should be used to estimate the reliability of speed tests. As no retest was carried out in the present study we decided to report Cronbach's Alpha, as it is reported in the test manual, too. According to the test manual (Brickenkamp, 2002), test-retest reliability is .74 up to .89 within a period of 3 months.

Gender differences and anonymity

Before the second and third hypotheses can be investigated, the influence that gender and anonymity had upon attention score was examined with a 2 x 2 - ANOVA. This yielded a significant main effect for the factor anonymity with anonymous conditions resulting in a higher test score than under the non-anonymous condition, $F(1,130) = 5.91$, $p = .016$, $\eta^2 = .043$. Gender had no significant influence upon attention score, and the interaction between gender and anonymity was not significant, either. The descriptive results are printed in table 2. There is a tendency towards a higher attention score for the females than for the males which does not reach significance due to a somewhat lower effect size than the factor anonymity, and less power because of the lower number of males.

In order to examine the second and third hypotheses, a 2 x 2 - ANCOVA was carried out. The dependent variable was self-estimation, the first factor was gender, and anonymity was the second factor. To take the difference in attention score under anonymous and under nominal conditions into consideration, attention score was included in the ANCOVA as a covariate. The descriptive results are shown in table 3 under the heading "self-estimate". A significant main effect regarding gender was revealed, with women's self-estimates being considerably lower than men's, $F(1,129) = 28.09$, $p = .001$, $\eta^2 = .179$. The estimated means were $M = 28.9$ for women and $M = 48.0$ for men. No significant effects were found regarding anonymity and interaction between anonymity and gender.

Hypotheses two and three can also be examined by comparing the estimated percentages with the percentages that were actually achieved. To do so, a 2 x 2 x 2 - ANOVA with the between-factors gender and anonymity and the within-factor estimated vs. actual test score

Table 2:
Descriptive statistics of the attention score

	Anonymous			Nominal			Total		
	n	M	SD	n	M	SD	n	M	SD
Female	47	204.0	38.7	53	193.1	28.4	100	198.3	33.9
Male	23	198.2	41.8	11	174.1	23.9	34	190.4	38.3
Total	70	202.1	39.5	64	189.8	28.4	134	196.3	35.1

Table 3:
Descriptive statistics of the self-estimated percentage and the attention score converted into percentage

		Anonymous			Non-anonymous		
		n	M	SD	n	M	SD
Female	Self-estimate	47	32.8	18.0	53	25.9	16.0
	Attention score	47	56.6	31.6	53	48.7	24.7
Male	Self-estimate	23	49.4	25.1	11	42.7	10.1
	Attention score	23	50.0	32.9	11	31.2	21.6

was calculated. Table 3 depicts the descriptive statistics of each cell. As already calculated further above with a t-test, the students estimated their performance considerably lower than was actually the case, $F(1,130) = 9.37$, $p = .003$, $\eta^2 = .067$. Moreover, women underestimated themselves to a greater extent than men did, as the interaction between estimated vs. actual test score and gender shows, $F(1,130) = 24.12$, $p = .000$, $\eta^2 = .156$. All other effects were not significant.

Discussion

Self-estimation of attention performance

The first hypothesis of this study stated that students are able to self-estimate their own attention score. The significant correlation of .31 between retrospective self-estimation in social comparison to fellow students and the actual score achieved in the test is in the same magnitude as that of intelligence (cf. also table 1 and the meta-analysis mentioned before). Therefore, the order and the distances between the student's estimates correspond with their test result. This finding is limited by the fact that mainly predictive ratings were employed in the studies on self-estimation of intelligence which, according to the findings in the meta-analysis (Mabe & West, 1982), are not quite as successful as the retrospective estimates employed in this study. One could speculate that students based their estimates primarily on the number of letters they processed, as it is easier to estimate than the number of mistakes. Because processing speed, measured by the amount of letters that have been processed in the fixed time, and attention score correlated highly at $r(133) = .91$, $p = .000$, processing speed is a reliable indicator for test performance.

However, the correlation between self-estimate and test performance does not give any information about whether students overestimate or underestimate their performance. Only a comparison of the estimated percentages with the percentages actually achieved shows that the students underestimate themselves distinctly. At first, this finding seems surprising in view of social psychological research findings as studies have found that people tend towards self-enhancement and have positive rather than negative illusions about themselves (Greenwald, 1980). According to Greenwald, these positive illusions and distortions are functional for the survival of the self. Based on an extensive literature review, Taylor and Brown (1988, 1994) conclude that moderately positive illusions are conducive to adaptation to the environment and to mental health, and that realistic estimates are rather typical of depressive persons. However, above all, this should be valid for "qualities being appraised that lack objective referents. The inherent subjective nature of such attributes means that individuals are more or less free to decide for themselves, how loyal, responsible, sincere, or trustworthy they are" (Brown, 1986, p. 371). However, Taylor and Brown's review has been criticized on account of renewed reviews of empirical findings and on account of their argumentation, and it has been argued that a more realistic self-estimate is conducive to good health, at least in the long term (Colvin & Block, 1994). From the standpoint of attribution theory, too, realistic estimates are functional for achieving personal (performance) goals (Försterling, 1994). A realistic estimate of one's own ability can save oneself from taking on tasks that can only be managed with considerable time and effort. Thus, more resources are available for more promising activities (Försterling & Morgenstern, 2002). The suggestion of

distinguishing situations in which positive illusions are beneficial from situations in which undistorted self-estimates are of use to a person (Robins & Beer, 2001) aims in a similar direction. In our study, the students knew they were going to receive objective feedback later. To that extent, negative illusions about one's own capability are one way of avoiding failure that is hard to attribute externally and uncontrollably when receiving feedback. Underestimating oneself raises the chances of doing better than expected. Therefore, negative illusions can be quite conducive to self-enhancement under the conditions given in this study. An alternative explanation could be that the participants felt they were underperforming, because the lines of the d2 are too long to process all the letters within the given 20 seconds per line. Thus, it is almost inevitable that a certain amount of letters will remain unprocessed, and this makes participants feel they have not done well which, in turn, leads to the underestimation of their own performance.

Gender differences

As stated in the second hypothesis, we expected men's self-estimates to be higher than those of women, even considering the actual attention score. This hypothesis can be retained, as the ANCOVA with attention score as a covariate revealed. Because actual performance was statistically controlled, factually superior performances can be ruled out as an explanation for higher male self-estimations, especially since the participating females achieved a higher attention score than the males, descriptively speaking. In a US study using the d2, this difference even became significant (Bates & Lemay, 2004), so it is more probable that women's attention scores are, in fact, superior, and that actual differences can not be the reason for men's higher self-estimates. The comparison between estimated and actual percentage also points in favour of retaining the hypothesis. The size of effect was in a small to medium range. This is in accordance with most of the other gender effects that were found (Eagly, 1995).

The results maintain that social stereotypes are responsible for the differences between male and female self-estimation, similarly to self-estimates of intelligence. As early as twenty years ago, gender researchers were demanding that social scientists should turn towards social stereotypes relating to gender to explain gender differences (Deaux, 1984; Deaux & LaFrance, 1998, p. 793ff). Accordingly, stereotypes consist of expectations concerning performance in typically male respectively typically female tasks, with expectations being higher when gender and task type match than when they do not. An alternative social psychological explanation referring to stereotypes concerning performance could be the stereotype that women want to portray themselves in a more modest light than men. If this were the case, it would have to result in an interaction effect not discovered in this study between gender and anonymity, because being motivated to present oneself modestly should not create an effect under anonymous conditions. In two similarly set up experiments by Heatherington et al. (1993) with self-estimated GPA (grade point average) scores, the interaction effect was not significant, either, but a significant gender effect under the anonymous condition (one-tailed t-test without alpha adjustment) was reported: When the non-anonymous condition prevailed, women estimated their GPA score on a lower level than men did. However, Heatherington et al. employed a stronger treatment in their experiments than the one employed in the present study to produce non-anonymity, namely by asking the

participants to give the examiner their estimates orally. Furthermore, the test's power was greater than in the present study because of the bigger sample, the use of single t-tests instead of an ANOVA, and the alpha adjustment not being carried out.

Influence of anonymity on self-estimation of attention

The third hypothesis postulated that self-estimates are more accurate when they are anonymous. This hypothesis can not be maintained in the present study, even considering that better test scores were achieved anonymously than nominally as a covariate in the ANCOVA. It is fair to suppose that solely invalidating the anonymity is not sufficient to activate the need to protect one's own self-esteem to a greater extent than under anonymous conditions. And what is more, the participants underestimated themselves distinctly, as discussed further above. As mentioned before, anticipation of feedback could have lead to a self-enhancing distortion under both circumstances. This was also suggested in Mabe and West's (1982) meta-analysis. A much higher effect size with a standardized beta-coefficient of .42 was reported there when an objective examination of the self-estimate was expected than under anonymous conditions, where the standardized beta-coefficient was .15. Possibly, the anticipation effect reduced the variance because of the relatively weak effect of the anonymity condition. In addition, the sample of 134 students was sufficient to test for a medium effect, but one would need 788 participants for a test of a small effect to become significant (Bortz & Döring, 2002, p. 616), as is to be expected under the anonymity condition.

Limitations of the study and conclusions

The low amount of males compared to the females in the present study is due to the fact that around 70 % of the teacher students at the university the data was collected at are females. Thus, the sample reflects the university population quite well. Nevertheless, the findings need to be replicated with populations and samples where females and males are equally distributed. Most important for future studies is the inclusion of other concentration and attention tests in order to investigate whether the results can be found for other operationalizations of attention, too. Futher, the thesis that anticipating feedback leads to negative illusions to avoid failure should be examined. This could be done by experimentally varying anticipation of feedback, for example. Furthermore, the explanation of gender differences with stereotypes could be examined with the aid of so-called male and female tasks. By means of a sample that is as large and representative as possible, the question whether attention is classified as typically male or not should be settled. Alternatively, a social psychological approach could be put into effect by experimentally modifying the instructions in a manner that would portray attention as typically male or typically female (cf. Deaux, 1984). Regarding anonymity, the experimental variation should be designed more intensely in order to activate both of the postulated needs for accurate estimation and protection of self-esteem more effectively. For example, in addition to the variation nominal vs. anonymous, it would be conceivable to vary consequences such as a cash prize for accurate estimates.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *The standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bates, M. E., & Lemay, E. P., Jr. (2004). The d2 test of attention: construct validity and extensions in scoring techniques. *Journal of the International Neuropsychological Society, 10*(3), 392-400.

Beloff, H. (1992). Mother, father and me: our IQ. *The Psychologist, 5*, 309-311.

Bennett, M. (1996). Men's and women's self-estimates of intelligence. *Journal of Social Psychology, 136*(3), 411-412.

Beyer, S., & Bowden, E. M. (1997). Gender differences in self-perception: convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin, 23*(2), 157-172.

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*(3), 546-553.

Bortz, J., & Döring, N. (2002). *Forschungsmethoden und Evaluation* (2 ed.). *[Research methods and evaluation* (2 ed.)*]* Berlin, Heidelberg: Springer.

Brickenkamp, R. (2002). *Test d2. Aufmerksamkeits-Belastungs-Test* (9 ed.). *[The d2 test. Test of attention under pressure* (9 ed.).*]* Göttingen: Hogrefe.

Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgements. *Social Cognition, 4*(4), 353-376.

Byrd, M., & Stacey, B. (1993). Bias in intelligence perception. *The Psychologist, 6*, 16.

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Colvin, C. R., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin, 116*(1), 3-20.

Cohen, J. (1988). *Statistical power analysis for the behavioural science*. New York: Erlbaum.

Deaux, K. (1984). From individual differences to social categories: analysis of a decade's research on gender. *American Psychologist, 39*(2), 105-116.

Deaux, K., & LaFrance, M. (1998). Gender. In D. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *Handbook of social psychology* (4 ed.). New York: Random House.

Eagly, A. H. (1995). The science and politics of comparing woman and men. *American Psychologist, 50*(3), 145-158.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117-140.

Fingerman, K. L., & Perlmutter, M. (1994). Self-ratings of past, present, and future cognitive performance across adulthood. *International Journal of Aging and Human Development, 38*(4), 363-382.

Försterling, F. (1994). The functional value of realistic causal attributions. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 5, pp. 151-179). Chichester, England: Wiley.

Försterling, F., & Morgenstern, M. (2002). Accuracy of self-assessment and task performance: Does it pay to know the truth? *Journal of Educational Psychology, 94*(3), 576-585.

Furnham, A., Clark, K., & Bailey, K. (1999). Sex differences in estimates of multiple intelligences. *European Journal of Personality, 13*, 247-259.

Furnham, A., Fong, G., & Martin, N. (1999). Sex and cross-cultural differences in the estimated multifaceted intelligence quotient score for self, parents and siblings. *Personality and Individual Differences, 26*, 1025-1034.

Furnham, A., & Rawles, R. (1999). Correlation between self-estimated and psychometrically measured IQ. *Journal of Social Psychology, 139*, 405-410.

Gardner, H. (1998). Are there additional intelligences? The case for naturalistic, spiritual, and existential intelligences. In J. Kane (Ed.), *Education, information, and transformation* (pp. 111-131). Englewood Cliffs, NJ: Prentice Hall.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35*(7), 603-618.

Heatherington, L., Daubman, K. A., Bates, C., Ahn, A., Brown, H., & Preston, C. (1993). Two investigations of "female modesty" in achievement situations. *Sex Roles, 29*(11/12), 739-754.

Hogan, H. W. (1978). IQ self-estimates of males and females. *Journal of Social Psychology, 106*(1), 137-138.

Holling, H., & Preckel, F. (2005). Self-estimates of intelligence - methodological approaches and gender differences. *Personality and Individual Differences, 38*(3), 503-517.

Jäger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen: Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells. [Multimodal classification of intelligence performance: experimentally controlled development of a descriptive intelligence structure model.] *Diagnostica, 28*(3), 195-225.

Jäger, A. O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven. [Intelligence structure research: rivalling models, new developments, prospects.] *Psychologische Rundschau, 35*(1), 21-35.

Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test. BIS-Test. Form 4. Handweisung. [Berlin intelligence structure test. BIS test. Form 4. Manual.]* Göttingen: Hogrefe.

Jonsson, A.-C., & Allwood, C. M. (2003). Stability and variability in the realism of confidence judgments over time, content domain, and gender. *Personality and Individual Differences, 34*, 559-574.

Lynn, R. (1998). Sex differences in intelligence: data from a Scottish standardisation of the WAIS-R. *Personality and Individual Differences, 24*(2), 289-290.

Lynn, R., & Irwing, P. (2002). Sex differences in general knowledge, semantic memory and reasoning ability. *British Journal of Psychology, 93*, 545-556.

Mabe, P. A., III, & West, S. G. (1982). Validity of self-evaluation of ability: a review and meta-analysis. *Journal of Applied Psychology, 67*(3), 280-296.

Masters, M. S., & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behaviour Genetics, 23*, 489-499.

Paulhus, D. L., Lysy, D. C., & Yik, M. S. M. (1998). Self-report measures of intelligence: Are they useful as proxy IQ tests? *Journal of Personality, 66*(4), 525-554.

Paulhus, D. L., & Morgan, K. L. (1997). Perceptions of intelligence in leaderless groups: The dynamic effects of shyness and acquaintance. *Journal of Personality and Social Psychology, 72*(3), 581-591.

Rammstedt, B., & Rammsayer, T. H. (2000). Sex differences in self-estimates of different aspects of intelligence. *Personality and Individual Differences, 29*, 869-880.

Rammstedt, B., & Rammsayer, T. H. (2002). Self-estimated intelligence: Gender differences, relationship to psychometric intelligence and moderating effects of level of education. *European Psychologist, 7*(4), 275-284.

Reilly, J., & Mulhern, G. (1995). Gender differences in self-estimated IQ: The need for care in interpreting group data. *Personality and Individual Differences, 18*(2), 189-192.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: short-term benefits and long-term costs. *Journal of Personality and Social Psychology, 80*(2), 340-352.

Schmidt-Atzert, L., & Bühner, M. (2000). Aufmerksamkeit und Intelligenz. [Attention and intelligence.] In *Intelligenz und Kognition*. Landau: Verlag Empirische Pädagogik.

Schmidt-Atzert, L., Bühner, M., & Enders, P. (2006). Messen Konzentrationstests Konzentration? Eine Analyse der Komponenten von Konzentrationsleistungen. [Do concentration tests measure concentration? An analysis of the components of concentration performance.] *Diagnostica, 52*(1), 33-44.

Scholz, A. (2006). *Konzentration im Alltag: Erleben und Leistung. [Concentration in everyday life: experience and achievement.]* Dissertation, Fakultät Mathematik und Naturwissenschaften der Technischen Universität Dresden, Dresden.

Schweizer, K., Moosbrugger, H., & Goldhammer, F. (2005). The structure of the relationship between attention and intelligence. *Intelligence, 33*(6), 589-611.

Stankov, L. (1998). Calibration curves, scatterplots and the distinction between general knowledge and perceptual tests. *Learning and Individual Differences, 8*(1), 29-50.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social-psychological perspective on mental health. *Psychological Bulletin, 103*(2), 193-210.

Taylor, S. E., & Brown, J. D. (1994). Positive illusions and well-being revisited: Separating fact from fiction. *Psychological Bulletin, 116*(1), 21-27.

Westhoff, K., & Hagemeister, C. (2005). *Konzentrationsdiagnostik. [Concentration diagnosis.]* Lengerich: Pabst.