

## Research Reports

# The Paradoxical Effect of Praise and Blame: Age-Related Differences

Kai Kaspar<sup>\*ab</sup>, Helena Stelz<sup>a</sup>

[a] Institute of Psychology, University of Osnabrück, Osnabrück, Germany. [b] Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany.

## Abstract

Blame and praise sometimes have a seemingly paradoxical effect: blame after failure sometimes leads to the impression that the recipient has a high ability. In contrast, praise after success can lead to an inference of low ability. Several studies showed this effect and its relation to concurrent sympathy ratings as well as to cognitive schemas underlying these estimations. The present study is the first to focus on differences across the adult lifespan regarding peoples' ability and sympathy estimations of praised or blamed individuals. In this context, the causal schemas underlying peoples' judgments were also investigated. The results show that seemingly paradoxical ability estimations were independent of age in contrast to reversed ability estimations (i.e., a praised individual is perceived as more capable than a blamed individual) decreasing with age. Moreover, age affected the estimation of a teacher's sympathy for blamed and praised students in terms of a universal decrease with age. Finally, exclusive ability-related causal schemas were primarily used by older people and facilitated the occurrence of seemingly paradoxical ability estimations. The results provide a first base for future research on this largely neglected topic.

*Keywords:* praise, blame, paradoxical ability estimations, causal schema, age differences

Europe's Journal of Psychology, 2013, Vol. 9(2), 304–318, doi:10.5964/ejop.v9i2.540

Received: 2012-10-23. Accepted: 2013-02-05. Published: 2013-05-31.

\*Corresponding author at: Institute of Psychology & Institute of Cognitive Science, University of Osnabrück, Seminarstr. 20, 49074 Osnabrück/Germany.  
E-mail: kkaspar@uos.de



This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Introduction

Evaluative behavior has an informative function since it provides data about how the ability of the recipient is perceived by the evaluating person (Meyer, 1982). When someone is blamed after failure in a performance setting (i.e., someone finds fault in somebody's performance), this can lead to an inference of high ability. In contrast, praise after success (in terms of expressing commendation for a good performance) can lead the recipient or an impartial observer to the impression that the ability of the recipient is estimated as low by the praising person. This effect might seem paradoxical at first sight and hence was labeled as the "paradoxical effect" on ability attribution that was shown by various studies (e.g., Barker & Graham, 1987; Binser & Försterling, 2004; Hofer, 1985; Meyer et al., 1979; Miller & Hom, 1996; Möller, 1999; Reisenzein, Debler, & Siemer, 1992). However, Möller (1999) pointed out that this effect is not actually paradoxical, rather it only appears paradoxical. Indeed, different models have been proposed which try to explain this effect:

### Attribution Model

On the basis of two principles from attribution theory, we can explain such seemingly paradoxical ability estimations: on the one hand, the so called "effort schema" postulated by Weiner and Kukla (1970) says that praise and blame

are commonly used when the outcome of an action is obviously derived from the effort of the acting person. According to this principle, blame will more likely be expressed if the outcome of an action is linked to low effort. On the other hand, according to the “compensatory principle” (Kukla, 1972) the evaluating person considers the degree of effort which is necessary for the specific outcome given a specific ability level. Hence, success will be attributed to a high ability if the evaluating person assumes that the acting person made less of an effort to reach the positive outcome. However, if the positive outcome is attributed to high effort, the ability of the acting person will be estimated as low. In contrast, a negative outcome will be ascribed to low ability if it arises despite of high effort but the acting person’s ability will be estimated as high if the evaluating person supposes low effort. Consequently, blame for failure can lead the recipient or an impartial observer to the conclusion that the blaming person estimates the ability of the blamed person as high but assumes low effort.

### Language-Psychological Model

Grice (1975) postulated four maxims jointly expressing the general cooperation principle of conversation (for a detailed discussion of the maxims see Schwarz, 1996). In short, according to this principle, the recipient of a message is entitled to expect that the speaker attempts to be informative, clear, relevant, and truthful, meaning that the speaker and the recipient communicate in an effective and economic way by means of cooperation. Based on this assumption, Groeben and Blickle (1988) derived an alternative explanation of paradoxical ability estimations: they argue that the cooperation principle is violated when a teacher gives different verbal feedbacks (praise and blame) to students who behave identically. The students (or external observers) should recognize this violation as intended by the teacher and consequently they should try to infer the meaning of the teacher’s statements by considering the current circumstances (see Blickle, 1993, p. 48). At the end, the students (or the external observers) may come to the conclusion that the teacher’s intention was to express an estimation of the recipients’ abilities. Hence, based on language psychology seemingly paradoxical ability estimations can be reconstructed, whereby the central mechanism is the active construction of sense from the speaker’s messages. In contrast, the attribution model described above does not postulate such a communicative intention by the speaker. Moreover, after Reisenzein (1990) had claimed that the language-psychological model of Groeben and Blickle (1988) would suffer from several substantial shortcomings and obscurities, an empirical study of Reisenzein et al. (1992) raised some doubts that this model is correct. They compared this model experimentally with the attribution model and only found support for the latter one. In reaction to this, Groeben and Blickle (1992) asserted that the study by Reisenzein et al. (1992) suffered from methodological problems. Reisenzein and Battmann (1992) in turn rejected this criticism and pointed out that no independent data supporting Groeben and Blickle’s model existed. Finally, Blickle (1993) softened the strict distinction between the attributional and language-psychological model by partially merging them (p. 45). Correspondingly, Reisenzein (1990) had previously proposed that the two models do not compete, but instead are rather integrable.

### Expectation Discrepancy Model

Hofer (1985) criticized the attribution model since he believes that no effort calculation (i.e., effort schema) would be necessary to explain seemingly paradoxical ability estimations. According to the attribution model, the students (or external observers) must believe that the teacher expresses praise and blame as he (or she) perceived differences in the students’ efforts. Otherwise, the second process (i.e., the compensatory principle) could not work. Hofer (1985) performed a study in which two students were treated differently by a teacher (blame vs. praise) after failure. In addition to this classical scenario, the study participants received the information that both students had made the same effort. Although the effort schema should not work in this case, Hofer found seemingly para-

doxical ability estimations. Therefore, he supposed an alternative interference according to which the recipient (or an external observer) thinks in the following way: “when the teacher believes a pupil is capable of more than another pupil he expects a better result. If the achieved result is lower than expected the teacher expresses his disappointment as a blaming utterance” (p. 415). In line with this, a student should be praised if the achieved result is higher than expected by the teacher. Consequently, this explanation model does not assume a mediating effort schema. Rather, it assumes a more direct mechanism producing seemingly paradoxical ability estimations: when giving feedback in a performance setting, it is assumed that the teacher takes into account his beliefs about the students’ abilities, and these beliefs directly determine the valence of the feedback.

To conclude, all three theoretical approaches describe a plausible mechanism that could produce seemingly paradoxical ability estimations, i.e., the teacher’s unequal treatment of the students is attributed to differences in the students’ ability. While the postulated processes differ, all three models converge to this point in the end. Nonetheless, the majority of studies in the area of praise and blame referred to the attribution model. It is the one that has been validated more extensively - also in the recent past (e.g., [Binser & Försterling, 2004](#); [Möller, 1999](#)). However, while we must note that there has been unsolved controversy about the mechanisms behind the seemingly paradoxical effect of praise and blame two decades ago, the existence of this effect is not questioned. Paradoxical ability estimations revealed in the context of written vignettes describing classroom scenarios (e.g., [Binser & Försterling, 2004](#)) were replicated with videotaped scenes ([Barker & Graham, 1987](#)) and were shown in experimental studies with a more realistic setting ([Meyer, 1982](#); [Meyer, Mittag, & Engler, 1986](#)). Moreover, a longitudinal study by [Tacke and Linder \(1981\)](#) found that students who were not praised by their teachers, in contrast to students who were praised for the same responses involving memory, improved their self-concept, whereby the improvement was domain-specific. Across twelve lessons non-praised students improved their self-estimated memory ability, but a generalization to other self-concepts such as creativity was not found.

[Binser and Försterling \(2004\)](#) analyzed several studies and reported a mean frequency of 64% of subjects showing seemingly paradoxical ability estimations, while the frequency differed substantially between studies. The occurrence of this paradoxical effect of praise and blame depends on several factors which were revealed in the context of the attribution model: Firstly, it is required that the praising and blaming person knows the ability of the blamed person ([Meyer & Plöger, 1979](#)). Secondly, the probability of the effect is strengthened in the context of difficult tasks, since effort (and hence the effort schema) does not play a substantial role in simple tasks ([Meyer, Reizenzein, & Dickhäuser, 2004](#)). Thirdly, paradoxical interpretations of teacher-student interactions are boosted if the evaluating person is asked to think about the teacher’s possible attributions and his/her ability estimation ([Rheinberg & Weich, 1988](#)). However, paradoxical interferences also occurred when ability was not considered as a possible reason for the teacher’s behavior ([Meyer, Bedau, & Engler, 1988](#)). Fourthly, it must be known that the teacher uses a normative reference system focussing on social comparisons, instead of an individual reference system for reward ([Möller, 1999](#)). Fifthly, the effect is more likely to occur if subjects use an ability-related causal schema for the interpretation of the teacher’s behavior. In contrast, the paradoxical effect is less likely to occur when subjects perceive the blame and praise in terms of sympathy ([Binser & Försterling, 2004](#); [Meyer et al., 2004](#)). Finally, the age of those individuals who evaluate the teacher-student interaction is a central factor for paradoxical ability estimations.

Previous studies considering the potential impact of age were limited to samples of children: results from a study by [Meyer \(1978\)](#) indicate huge differences between younger and older children regarding the frequency of seemingly paradoxical ability ratings. [León-Villagrà, Meyer, and Engler \(1990\)](#) systematically compared children

of different ages and found that children under 9 years do not show seemingly paradoxical ability estimations. [Barker and Graham \(1987\)](#) suggest that children do not show such effects because they do not apply the compensatory principle. Finally, [Rheinberg, and Weich \(1988\)](#) showed that the frequency of this seemingly paradoxical effect increases linearly with age when only considering a range from 13 to 18 years. However, it is still unclear whether this linear trend remains across the whole lifespan, whether it asymptotically approaches an upper limit, or if it even decreases at a certain age.

### The Present Study

The present study was conducted in order to investigate the effect of seemingly paradoxical ability estimations across the whole lifespan. It is important to note that previous literature has completely neglected potential effects of age cohorts in the area of praise and blame. Against this background, the present results would be of importance for our understanding of the mechanisms behind this seemingly paradoxical phenomenon. If the frequency of this effect increased across the lifespan, we would have to rethink the tacit idea (shared by the different theoretical approaches) that an age-invariant cognitive mechanism (with the exception of young children) mediates the seemingly paradoxical effect of praise and blame across the whole lifespan. Although the level of cognitive development should not be the moderating factor in adult age, previous findings suggest age-related differences in social judgments across the adult lifespan (e.g., [Blanchard-Fields, 1994, 1996](#); [Hess & Follett, 1994](#)). Against the background of the attributional model it is important to note that older adults were found to use different schemas to explain social behavior ([Blanchard-Fields, 1996](#)). All three explanation models described above focus on ability differences as the cause constituting seemingly paradoxical ability estimations. [Binser and Försterling \(2004\)](#) broadened the view by also considering potential differences in the teacher's sympathy for the students as responsible for the teacher's praising or blaming remarks. In the present study, we will consider both potential explanations, hereafter referred to as "causal schemas".

In addition to the usage of different causal schemas, the degree and the quality of judgments are influenced by the type and content of social dilemmas presented to study participants: [Blanchard-Fields \(1996\)](#) found the variability in age differences in attributional responding as a function of the specific content domain in question. She demonstrated age and generational differences in the frequency of different schemas which were produced for different social dilemmas. For the present study this means we cannot simply derive conclusions from other research areas. No previous study addressed potential effects of age cohorts on the evaluation of the prototypical scenario that produces seemingly paradoxical ability estimations: a teacher treats students who behave identically unequally. Hence, we currently have very little knowledge of the influence of age in the context of this seemingly paradoxical effect. Therefore, the present study was conducted to scrutinize potential age effects on ability estimations by an external observer (in this case our study participants) when confronted with the classical classroom scenario ([Meyer et al., 1979](#)). In this scenario one student (A) is praised by a teacher whereas another student (B) is blamed for equal test performances.

As specific causal schemas might lead to different results in ability estimations, study participants' (i.e., the external observers') explanations for the teacher's unequal treatment of the students was considered as well as a potential moderating effect of the observer's age. In this context, [Binser and Försterling \(2004\)](#) showed that sympathy related causal schemas reduced the probability of paradoxical ability estimations, whereas differences in sympathy estimations increased. Furthermore, in a study by [Hofer and Pikowsky \(1988\)](#) adults reported more often than adolescents that the teacher's unequal treatment of the students was based on differences in sympathy. Hence, the

teacher's sympathy for both students from the perspective of an external observer was also considered. The procedure was closely modeled on previous studies to allow for comparisons.

The hypotheses in the present study are:

1. In accordance with [Binser and Försterling \(2004\)](#), three subgroups are expected which differ in the direction of the estimated ability differences (ability rating groups:  $A < B$  (seemingly paradoxical effect),  $A = B$ ,  $A > B$ ). In addition to Binser and Försterling, we asked whether the strength of the corresponding ability ratings is constant across the entire lifespan.
2. The direction of the estimated ability difference is expected to be independent of the evaluator's age.
3. The praised student is expected to receive higher sympathy ratings in all ability rating groups (as suggested by [Binser and Försterling, 2004](#)), whereby the amount of difference is expected to vary between the ability rating groups: the difference should be maximal in the group " $A > B$ ". Furthermore, we investigate whether the strength of estimated sympathy differences is independent of the evaluator's age within each ability rating group.
4. Subjects' explanations (i.e., causal schemas) for the teacher's unequal treatment of the students will trigger the signature of the subsequent ability and sympathy ratings as previously shown (e.g. [Binser & Försterling, 2004](#); [Meyer et al., 2004](#)). In this context, we investigated whether the evaluator's age influences the probability of a certain causal schema. Results of a study by [Hofer and Pikowsky \(1988\)](#) suggest that differences in the teacher's sympathy for both students would become a more prominent causal schema with increasing age of the evaluator.

## Method

### Participants

141 subjects with a mean age of 41.77 years ( $SD = 27.57$ ; range: 11 - 96) participated. Children under the age of 11 years were not included as they are not able to show paradoxical ability estimations ([León-Villagr a et al., 1990](#)). We used a German sample and all subjects were native speaker so that they were able to read and understand the written vignette describing the classroom scenario. Children were acquired via a convenient sample in two schools; young adults were recruited on the university campus. We got access to middle aged adults via snowball sampling. Older adults and seniors were acquired from a senior sports club. In order to get access to very old subjects we visited two retirement homes. Prior to the study, the home administration helped us to find a pre-selection of appropriate subjects (i.e., seniors with clinical mental disorders were excluded to prevent data biases).

The study conformed with the Code of Ethics of the American Psychological Association, to the Declaration of Helsinki, and to national guidelines. The study was initiated and coordinated entirely by the authors.

### Procedure and Materials

We used a standardized procedure for all participants. They all voluntarily participated in the study and did not receive incentives. At the beginning of a session, participants were introduced to the topic of the study, but we did not explain the actual purpose of the study at that time. Then the participants provided some demographic information. Afterwards, they were asked to carefully read a vignette describing the classical classroom situation in which a teacher treated two students who behaved identically unequally (e.g., [Binser & F orsterling, 2004](#); [Meyer et al., 1979](#)): two students successfully passed an easy test, but the teacher praised only one of them (student A), whereas the second student (B) was treated neutrally. Both students failed a difficult test, whereby only student B was blamed. Student A was treated neutrally despite failure. As [Groeben and Blickle \(1988\)](#) discussed

elaborately, sanctioning statements of a speaker (e.g., “good job” or “oh, not so great”) may be interpreted in different ways. Also, [Hofer \(1985\)](#) emphasized the ambiguity of verbal utterances, especially the potential divergence between the speaker’s actual intention to praise or blame on the one hand, and the recipient’s perception of the verbal utterances as praise or blame (or something else) on the other hand. For example, sarcastic communication can lead to quite different understandings of verbal utterances as illustrated in terms of “blame by praise” and “praise by blame” ([Anolli, Ciceri, & Infantino, 2002](#)). Hence, the recipient does not necessarily infer that specific utterances are signatures of praise or blame, respectively. In order to overcome this interpretation problem, we created a scenario in which we directly spoke of a teacher who is praising and blaming students (in German the corresponding verbs “loben” and “tadeln” are very unambiguous). In order to make the scenario more comprehensible, especially for the young participants, we introduced line-drawing figures depicting the conditions as done by [León-Villagr a et al. \(1990\)](#). After they had read the scenario, participants were asked to report the reason that could explain the teacher’s unequal treatment of the students. These causal schemas underlying the observers’ subsequent judgments were investigated by open-ended questions (cf. [Binser & F rsterling, 2004](#); [Blickle, 1991](#); [M ller 1999](#); [Rheinberg & Weich, 1988](#)). This is important to mention because a closed item format may prime specific ability estimations. Finally, participants answered two questions by rating on a scale from 1 (“very low”) to 9 (“very high”): “How high does the teacher assess the ability of student A/B?” and “How high is the teacher’s sympathy for student A/B?” (these are approximate translations of the original German wording). This single item format was selected in accordance with previous studies (e.g., [Binser & F rsterling, 2004](#); [Blickle, 1991](#); [M ller, 1999](#)), and the 9-point scale was chosen to compare absolute rating values in the present study with those found by [Binser and F rsterling \(2004\)](#). A small pre-test with some children showed, however, that it was mandatory to change wording for them because they had some problems to correctly understand the words “ability” and “sympathy”. We substituted these words and asked instead how smart/clever/intelligent does the teacher think the students are, and how much the teacher would like the students, respectively (the German wording we used was unambiguous). The sequence of ability and sympathy rating was counterbalanced across participants to prevent sequence effects.

## Data Analysis

The reported explanations for the teacher’s unequal treatment of the students were categorized following a two-step procedure ([Kaspar, Hamborg, Sackmann, & Hesselmann, 2010](#); [Kaspar & K nig, 2011](#)). This procedure first includes a categorization by two independent raters with the aid of a given category-system consisting of four categories:

- “ability”: the teacher’s unequal treatment of students is attributed to ability: participants assume that the teacher treats the students unequally because they differ in their general ability.
- “sympathy”: participants’ statements address sympathy but no ability aspects.
- “ability/sympathy”: participants reported ability-related reasons as well as sympathy reasons.
- “residual”: participants reported other reasons such as selective discrimination, pedagogical inability or more exotic explanations.

The categories were selected following [Binser and F rsterling \(2004\)](#) who, however, used an additional category for explanations which did not differentiate between ability and previous performance. This category did not provide a more specific result pattern but may have lowered the discriminatory power of the entire category system. The inter-rater-reliability in the present study was very high (*Cohen’s Kappa* = .95) and in the few cases of absent agreement a consensual categorization was forced in the second step to allow frequency analyses.

In order to statistically test our hypotheses, we focused on the ability and sympathy ratings and additionally took participants' ages as well as the causal schemas into consideration. For this, a moderated regression analysis and variance analytic approaches were used. Whenever important statistical assumptions were not fulfilled, reduced regression models or non-parametrical tests were applied. Initially, we calculated difference values with respect to the ability and sympathy rating, respectively. The ability rating for the blamed student B (*"How high does the teacher assess the ability of student B"*) was subtracted from the corresponding rating for the praised student A. The same difference was calculated for the sympathy ratings (*"How high is the teacher's sympathy for student A/B"*).

## Results

### Ability Ratings (Hypotheses 1, 2)

In accordance with Binser and Försterling (2004) we found three subgroups differing in the signature of their ability ratings. A substantial subsample of 59 participants (43.2%) rated student B's ability higher than student A's ability (ability rating group:  $A < B$ ), therefore showing the seemingly paradoxical effect. A second group of 65 subjects (46.8%) showed the reverse effect ( $A > B$ ), and a third small group of 14 subjects (10.1%) showed no difference in the ability ratings ( $A = B$ ). Hence, the ability rating group was introduced as an additional factor in subsequent analyses.

In order to clarify whether the strength of the corresponding ability ratings was constant across the whole lifespan, a moderated regression analyses was computed. The difference in the ability ratings (A minus B) served as criterion. The evaluator's (i.e., participant's) age was treated as a continuous predictor and the ability rating groups as a moderator in terms of dummy variables. Each dummy coded level of the moderator was additionally crossed with age to test for a potential interaction between rating groups and age. For this, group " $A > B$ " was selected as the reference group. In addition, the analysis was done without the ability rating group " $A = B$ " due to problems of multicollinearity ( $VIF > 10$ ).

Both multiple regression analyses (both  $R = 922$ ,  $p < .001$ ) revealed a significant predictive value of the evaluators' age as depicted in Table 1. The dummy variable for the group showing a paradox ability estimation " $A < B$ " had a significant predictive value which was moderated by participants' age. The group showing the reverse ability estimation " $A > B$ " (reference group) differed by about eight rating points. The significant interaction term indicates that the slope of the ability rating on age in group " $A < B$ " was .022 rating points greater than the slope of the ability rating on age in the reference group " $A > B$ ". In order to illustrate this differential effect of age on ability estimations, bivariate regressions were computed separately for both groups (seemingly paradoxical effect " $A < B$ ":  $R = .070$ ,  $B = .004$ ,  $p = .596$ ; non-paradoxical effect " $A > B$ ":  $R = -.276$ ,  $B = -.018$ ,  $p = .026$ ) and depicted by scatter plots (Figure 1, left side). Also, rating group " $A = B$ " differed significantly from group " $A > B$ ".

In summary, all groups differed significantly in their ability ratings, whereby the value of the seemingly paradoxical rating effect ( $A < B$ ) remained constant across the whole lifespan in contrast to the value of the reversed effect ( $A > B$ ) that decreased with age.

Table 1

## Result of the Multiple Regression Analysis

Variable	Ability rating			Sympathy rating		
	b	t	p	b	t	p
Age	-.018	-2.228	.028*	-.026	-2.016	.046*
Ability rating group "A < B"	-8.135	-16.121	<.001***	-1.997	-2.434	.016*
Age x Ability rating group "A < B"	.022	1.978	.050*	-.008	-.441	.660
Ability rating group "A = B"	-4.227	-2.781	.006**	2.105	.830	.408
Age x Ability rating group "A = B"	.018	.842	.402	-.051	-1.441	.152
<b>Reduced regression analysis without group "A = B"</b>						
Age		-2.337	.021*		-2.001	.047*
Ability rating group "A < B"		-16.907	<.001***		-2.416	.017*
Age x Ability rating group "A < B"		-2.074	.040*		-.438	.662

Note. Results of the multiple regression testing the moderating effect of the ability rating group (A < B, A = B, and A > B) on participants' age regarding the ability rating and the sympathy rating. The moderator "rating group" was dummy coded, whereby group "A > B" served as the reference group. The t and p values for the reduced regression analysis (without group "A = B") are also depicted.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

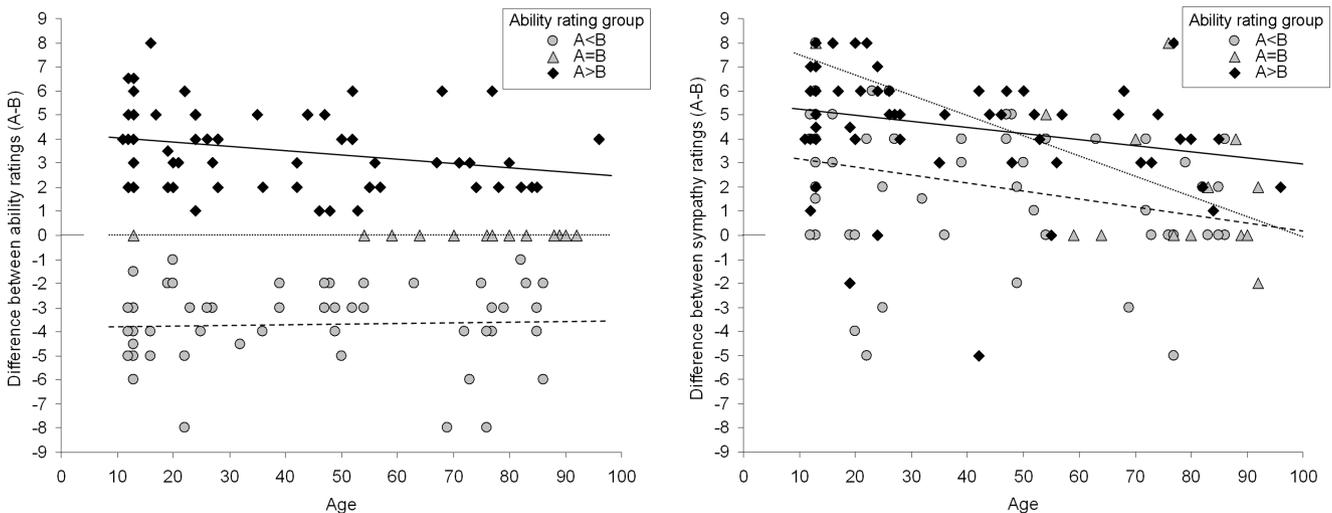


Figure 1. Scatter plots of participants' ability and sympathy ratings depending on age, and depicted regarding the three ability rating groups (A < B, A = B, A > B).

Note. The difference between the ability ratings for students A and B is shown on the left side, the rating difference in the teacher's sympathy for students A and B on the right side. Regression lines for all three ability rating groups are marked (dashed line: A < B; dotted line: A = B; solid line: A > B).

According to hypothesis 2, the direction of the rating effect was expected to be independent of the evaluator's age. However, this was not the case: the mean age differed between the rating groups (Kruskal-Wallis test:  $\chi^2(2) = 19.385, p < .001$ ). Post-hoc pairwise comparisons by means of Mann-Whitney U tests showed no difference in age between groups "A < B" ( $M = 40.66, SD = 26.86$ ) and "A > B" ( $M = 35.20, SD = 24.89; p = .222$ ). However, both groups were younger on average than the ability rating group "A = B" ( $M = 73.35, SD = 21.32$ ; both  $p < .001$ ).

### Sympathy Ratings (Hypothesis 3)

Regarding the sympathy ratings, the same moderated regression analysis as above was computed:  $R = 515$ ,  $p < .001$ . Again, the regression analysis was also done without the ability rating group “A = B” due to problems of multicollinearity:  $R = 496$ ,  $p < .001$ . Table 1 depicts the results. As expected, higher sympathy ratings were observed for the praised student in all ability rating groups but the difference between both students decreased with participants’ age. This decrease was not moderated by the ability rating group. Moreover, the difference in sympathy estimation was maximal in the ability rating group “A > B” which differed significantly from the group showing the paradoxical ability estimation (A < B). Figure 1 (right side) depicts the corresponding scatter plots.

### Analysis of Causal Schemas (Hypothesis 4)

Finally, we focused on the reported explanations for the unequal treatment of the students by the teacher. In order to clarify whether these causal schemas primed the subsequent ability and sympathy ratings in a specific way, we analyzed the ratings regarding the four categories used to classify the schemas. With respect to the ability ratings, differences between categories were present (Kruskal-Wallis test:  $\chi^2(3) = 12.779$ ;  $p = .005$ ) (Figure 2, left side). Post-hoc pairwise comparisons of schema categories by means of U-tests revealed a significant difference between those participants reporting ability reasons for the teacher’s behavior and participants reporting another reason (all  $U \geq 288.500$ ; all  $p \leq .034$ ).

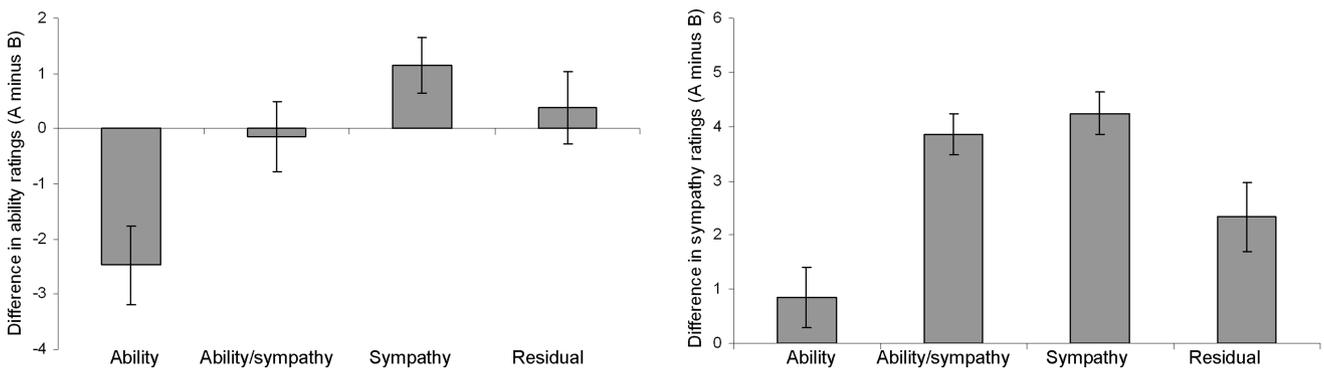


Figure 2. Participants’ ratings of the students’ ability (left side) and the teacher’s sympathy for the students (right side) depending on participants’ previously reported reason for the teacher’s behavior (praising student A, but blaming student B).

Note: Categories: “ability” = participants’ statements addressed ability aspects but no sympathy aspects; “ability/sympathy” = sympathy and ability aspects were reported; “sympathy” = statements addressed sympathy but no ability aspects; “residual” = an absence of any ability or sympathy aspects. Each participant was assigned to one of the four categories. Significant group differences are drawn in ( $\alpha_{adj} = .008$ ). Vertical lines above bars indicate standard error of the mean.

Regarding sympathy ratings we also found differences between subjects using different causal schemas (Kruskal-Wallis-test:  $\chi^2(3) = 26.517$ ;  $p < .001$ ) (Figure 2, right side). When only ability reasons were reported, the non-significant difference in sympathy ratings ( $t(19) = 1.522$ ;  $p = .145$ ) was significantly lower than when sympathy reasons were mentioned additionally ( $U = 144,500$ ; all  $p < .001$ ), or when sympathy reasons were reported exclusively ( $U = 173,000$ ; all  $p < .001$ ). Hence, when sympathy reasons were mentioned the teacher’s sympathy for the praised student A was rated as greater than his sympathy for the blamed student B. Moreover, participants who mentioned neither ability nor sympathy reasons (group “residual”) showed a significant difference in sympathy ratings ( $t(23) = 3.685$ ;  $p = .001$ ) which in turn was significantly smaller than in the two groups of participants mentioning sympathy reasons (both  $U \geq 329.500$ ; both  $p \leq .025$ ).

In summary, an ability-related causal schema facilitated the occurrence of seemingly paradoxical ability estimations but diminished differences in perceived sympathy. The reversed ability rating was shown by those participants reporting sympathy but no ability reasons.

Moreover, we found differences in age between categories (Kruskal-Wallis-test:  $\chi^2(3) = 22.381$ ;  $p < .001$ ) which were compared in pairs using U-tests. Participants mentioning only sympathy reasons (mean age = 38.24,  $SD = 25.54$ ) and those who additionally reported ability reasons ( $M = 29.27$ ,  $SD = 21.64$ ) differed nearly significantly ( $Z = -1.960$ ,  $p = .050$ ). Both groups were younger than participants who exclusively reported ability reasons ( $M = 53.55$ ,  $SD = 26.86$ ) (both  $Z \geq -2.125$ , both  $p \leq .034$ ) and also younger than those participants who did not mention any ability or sympathy aspects ( $M = 59.62$ ,  $SD = 29.07$ ) (both  $Z \geq -2.826$ , both  $p \leq .005$ ). Finally, subjects belonging to the residual category group did not differ in age from those mentioning exclusive ability-related explanations ( $Z = -.976$ ,  $p = .329$ ).

In order to ascertain whether seemingly paradoxical ability estimations necessarily followed whenever participants' ratings were based on an ability-related causal schema, we statistically compared the number of participants showing the seemingly paradoxical ability rating ( $A < B$ ) with the number of participants showing the reversed ability rating ( $A > B$ ) in each statement category by means of Chi squared tests. Participants using an ability-related causal schema showed the seemingly paradoxical ability rating more often than not ( $n = 15$  vs.  $n = 4$ ;  $\chi^2(1) = 6.368$ ,  $p = .012$ ). Participants assigned to the sympathy category showed the reversed ability rating more often than not ( $n = 17$  vs.  $n = 32$ ;  $\chi^2(1) = 4.592$ ,  $p = .032$ ). Regarding the other two categories, we did not find any differences (both  $\chi^2(1) \leq 1.471$ , both  $p \geq .225$ ). Consequently, the occurrence of seemingly paradoxical ability estimations was primed by previously mentioned ability reasons which were sometimes, however, followed by non-paradoxical ratings. Moreover, the seemingly paradoxical ability rating also occurred even when participants did not mention any ability aspects as relevant for the unequal treatment of students.

## Discussion

We found large differences in ability ratings in both directions on the level of subgroups (ability rating groups) as suggested by previous studies (e.g., [Binser & Försterling, 2004](#); [Meyer et al., 2004](#); [Möller, 1999](#)). Thereby, the value of seemingly paradoxical ability ratings ( $A < B$ ) remained constant across the whole lifespan, whereas the value of the reversed effect ( $A > B$ ) decreased with age (hypothesis 1). Hence, the linear increase of the frequency of this seemingly paradoxical effect found by [Rheinberg and Weich \(1988\)](#) did not continue across the adult age. Moreover, the probability of both effect directions was independent of the evaluator's age but a small third group estimating no ability difference between the students ( $A = B$ ) was about 35 years older on average (hypothesis 2). Consequently, the occurrence of seemingly paradoxical ability ratings did not depend on the age of adults, except very young children as previously shown ([Barker and Graham, 1987](#); [León-Villagr a et al., 1990](#)).

Interestingly, we found an overall trend with respect to sympathy ratings: in accordance with hypothesis 3, the teacher's sympathy for the praised student was rated significantly higher than his sympathy for the blamed student in all ability rating groups, but the difference between both students decreased with the study participants' age. This decrease was not moderated by the ability rating group, but the difference in sympathy estimation was maximal in the group " $A > B$ " as previously found by [Binser and F rsterling \(2004\)](#).

With respect to hypothesis 4, an ability-related causal schema facilitated the occurrence of subsequent paradoxical ability estimations ( $A < B$ ), but it diminished differences in perceived sympathy. Participants who reported

sympathy but not ability reasons for the teacher's unequal treatment of the students showed the opposite direction in ability estimation ( $A > B$ ) – on average. This non-paradoxical ability estimation was paralleled by larger differences between the teacher's sympathy for both students. On average, the teacher's sympathy for the praised student A was rated higher than his/her sympathy for the blamed student B, independent of the used causal schema. Moreover, participants who expressed a sympathy related causal schema were younger on average than participants who mentioned other explanations for the teacher's behavior. This result pattern contradicts what [Hofer and Pikowsky \(1988\)](#) found over two decades ago. In their study, adults in contrast to adolescents reported more often that the teacher's unequal treatment of the students would base on differences in sympathy. The present results suggest that older individuals are more ability-oriented when interpreting teacher-student interactions. In fact, age-related differences in schemas underlying social judgments have been found previously regarding several social situations ([Blanchard-Fields, 1996](#)).

All in all, the occurrence of seemingly paradoxical ability estimations was significantly primed by previously mentioned ability reasons. However, the occurrence of the reversed effect was also found on the base of ability-related schemas. Moreover, seemingly paradoxical ability ratings also occurred even when participants did not mention any ability aspects as relevant for the teacher's unequal treatment of the students. This result pattern suggests that seemingly paradoxical ability estimations are not completely determined by the causal schema used to explain the teacher's behavior. This result does not favor one of the above-described explanation models, namely the attribution model, the model of language-psychological sense construction, or the expectancy discrepancy model. As outlined, in all three theoretical models, the teacher's unequal treatment of the students is finally ascribed to the teacher's belief that one student is more capable than the other. It should also be noted that we can exclude potential differences in the participants' interpretation of the teacher's utterances (the central point in the language-psychological model), because we did not use specific verbal utterances to describe praise and blame. Rather, we relieved our study participants from the interpretation process as we created a scenario in which we directly spoke of a teacher who is praising and blaming students. Consequently, the cognitive mechanisms underlying paradoxical effects of praise and blame seem to be multi-factorial. The present results contradict a simple priming mechanism. An ability-related causal schema does not necessarily lead to seemingly paradoxical ability estimations but their value remained constant across the whole lifespan. Against this background, we need a revision of the current explanation models. From our point of view, each of them has its value but none of them provide reliable predictions. Perhaps it will be fruitful to create a more complex theoretical model instead to strive for less complexity as [Hofer \(1985\)](#) did with his expectation discrepancy model. However, before we will be able to formulate an elaborated model, future research should systematically compare the existing models on the one hand, since there has been only little endeavour to systematically compare the models so far. On the other hand, we need much more insight in the constraints which determine the occurrence of seemingly paradoxical ability estimations.

Overall, the presented data show that seemingly paradoxical ability estimations occur in all age groups, whereby the value of the seemingly paradoxical ability estimation remains constant across the entire lifespan. Thereby, older people show a preference for ability-related explanations when interpreting evaluative behavior. At the same time, sympathy differences between recipients of blame and praise, respectively, become less important with age. Hence, the evaluating person (e.g., a teacher, a superior, or a teammate) should carefully weigh the use of praise and blame in performance settings. This becomes especially important when a normative reference system is used focussing on social comparisons instead of an individual reference system for reward ([Möller, 1999](#)). Indeed, we should remember that the phenomenon of seemingly paradoxical ability estimations is not only an interesting effect that is worth to be addressed by basic research. From our point of view, it is even more important to consider

the motivational consequences which, for example, arise from the impression that a praising person believes that the recipient has a low ability in a certain domain.

In this context, further studies should also scrutinize whether the presented results can be generalized to other social interactions beyond the classroom scenario. Although most studies in the area of praise and blame have focused on teacher-student interactions in classroom scenarios so far, the results and implications can probably be generalized to other contexts in which performance and achievement motivation play a central role. Therefore, it is also important to further scrutinize the specific ability dimensions which are sensitive to seemingly paradoxical effects of praise and blame. Hofer (1985) did not find paradoxical effects with regard to students' ability to concentrate, their diligence, or their forgetfulness. The majority of the previous studies focused on (partially unspecific) task performances.

Furthermore, it would be interesting to see whether the seemingly paradoxical effect of praise and blame is culture-specific, or whether the causal schemas used to explain this phenomenon depend on cultural features. Perhaps we would find interactions between cultural specifics and certain ability domains.

Finally, we want to point out that most of the previous literature on praise and blame exclusively focused on verbal utterances. However, non-verbal cues also play a crucial role in communication, especially when interpreting ambiguous verbal messages. Non-verbal cues are not limited to facial expressions and gestures, but also include bodily sensations. Our evaluation of others significantly depend on incidental bodily sensations such as weight and texture (Ackerman, Nocera, & Bargh, 2010; Kaspar & Krull, 2013), and on incidental bodily interactions with the environment: for example, Kaspar (2013) recently showed that washing one's hands after failure in a performance setting enhanced optimism but hampered future performance in the same task domain. Perhaps, such incidental sensations also affect the way in which praise and blame are expressed and perceived. It could be very fruitful if research on praise and blame considered such variables to reveal the potential constraints of seemingly paradoxical effects and to build up a comprehensive theory. In the present study we presented a scenario description to our participants so that we exclude concrete verbal utterances and body language of the protagonists (except very rough line drawings). Thus, we probably reduced the inter-subject variance in the understanding of the scenario, but we also reduced the ecological validity of the situation at the same time. Under real-life circumstances any interaction between the sender of a message and its recipient is a multi-level process. This is important to consider with respect to the practical implications of the present research and regarding the design of future studies.

## References

- Ackerman, J. M., Nocera, C. C., & Bargh, J. A. (2010). Incidental haptic sensations of social judgments and decisions. *Science*, 328, 1712-1715. doi:10.1126/science.1189993
- Anolli, L., Ciceri, R., & Infantino, M. G. (2002). From "blame by praise" to "praise by blame": Analysis of vocal patterns in ironic communication. *International Journal of Psychology*, 37, 266-276. doi:10.1080/00207590244000106
- Barker, G. P., & Graham, S. (1987). Developmental study of praise and blame as attributional cues. *Journal of Educational Psychology*, 79, 62-66. doi:10.1037/0022-0663.79.1.62

- Binsler, M. J., & Försterling, F. (2004). Paradoxe Auswirkungen von Lob und Tadel: Personale und situative Moderatoren [Paradoxical effects of praise and blame: Personal and situative moderators]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 36, 182-189. doi:10.1026/0049-8637.36.4.182
- Blanchard-Fields, F. (1994). Age differences in causal attributions from an adult developmental perspective. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 49, P43-P51.
- Blanchard-Fields, F. (1996). Causal attributions across the adult life span: The influence of social schemas, life context, and domain specificity. *Applied Cognitive Psychology*, 10, 137-146.  
doi:10.1002/(SICI)1099-0720(199611)10:7<137::AID-ACP431>3.0.CO;2-Z
- Blickle, G. (1991). Anregungsbedingungen für scheinbar paradox(al)e Wirkungen von Lob und Tadel [Stimulatory conditions of apparently paradoxical effects of praise and blame]. *Zeitschrift für Pädagogische Psychologie*, 5, 21-32.
- Blickle, G. (1993). Attribution durch Konstruktion? Eine sprachpsychologische und attributionstheoretische Erklärung der Verarbeitung und Wirkung von Lob [Attribution by construction? A psycholinguistic and attributional explanation of the processing and of the effects of praise]. *Zeitschrift für Psychologie*, 201, 45-57.
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 4, pp. 41-58). New York: Academic.
- Groeben, N., & Blickle, G. (1988). Gegen einen objektivistisch halbierten Kognitivismus: Kognitiv-konstruktives Sprachverstehen und nicht-paradoxe Wirkungen von Lob und Tadel [Against cognitivism halved by objectivism cognitive-constructive understanding of languages and nonparadox effects of praise and blame]. *Zeitschrift für Sozialpsychologie*, 19, 85-102.
- Groeben, N., & Blickle, G. (1992). Anmerkungen zum "Verstehensvorgang bei scheinbar paradoxen Wirkungen von Lob und Tadel" nach Reizenzein, Debler und Siemer (1992) [Notes on „The interpretation process in the seemingly paradoxical effects of praise and criticism” by Reizenzein, Debler, and Siemer (1992)]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39, 684-688.
- Hess, T. M., & Follett, K. J. (1994). Adult age differences in the use of schematic and episodic information in making social judgments. *Aging and Cognition*, 1, 54-66. doi:10.1080/09289919408251450
- Hofer, M. (1985). Zu den Wirkungen von Lob und Tadel [About effects of praise and blame]. *Bildung und Erziehung*, 38, 415-426.
- Hofer, M., & Pikowsky, B. (1988). Wie Jugendliche bei freier Antwortmöglichkeit Lehrersanktionen deuten. *Zeitschrift für Pädagogische Psychologie*, 2, 243-250.
- Kaspar, K. (2013). Washing one's hands after failure enhances optimism but hampers future performance. *Social Psychological and Personality Science*, 4, 69-73. doi:10.1177/1948550612443267
- Kaspar, K., Hamborg, K.-C., Sackmann, T., & Hesselmann, J. (2010). The effectiveness of formative evaluation in the development of usable software: A case study. *Zeitschrift für Arbeits- und Organisationspsychologie*, 54, 29-38.  
doi:10.1026/0932-4089/a000003
- Kaspar, K., & König, P. (2011). Overt attention and context factors: The Impact of repeated presentation, image type and individual motivation. *PLoS ONE*, 6(7), e21719. doi:10.1371/journal.pone.0021719

- Kaspar, K., & Krull, J. (2013). Incidental haptic stimulation in the context of flirt behavior. *Journal of Nonverbal Behavior*. Advance online publication. doi:10.1007/s10919-013-0154-0
- Kukla, A. (1972). Foundations of an attributional theory of performance. *Psychological Review*, 79, 454-470. doi:10.1037/h0033494
- León-Villagrà, J., Meyer, W.-U., & Engler, U. (1990). Fähigkeitsinferenzen aufgrund von Lob: Eine entwicklungspsychologische Studie [Ability interference due to praise: A developmental psychological perspective]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 22, 54-74.
- Meyer, W.-U. (1978). Der Einfluß von Sanktionen auf die Begabungsperezeption [The impact of blame on ability perception]. In D. Görnitz, W.-U. Meyer, & B. Weiner (Eds.), *Bielefelder Symposium über Attribution* (pp. 71-78). Stuttgart: Klett.
- Meyer, W.-U. (1982). Indirect communications about perceived ability estimates. *Journal of Educational Psychology*, 74, 888-897. doi:10.1037/0022-0663.74.6.888
- Meyer, W.-U., Bachmann, M., Biermann, U., Hempelmann, M., Plöger, F.-O., & Spiller, H. (1979). The informational value of evaluative behavior: Influences of praise and blame on perceptions of ability. *Journal of Educational Psychology*, 71, 259-268. doi:10.1037/0022-0663.71.2.259
- Meyer, W.-U., Bedau, U., & Engler, U. (1988). Indirekte Mitteilungen über Fähigkeitseinschätzungen in hypothetischen Lehrer-Schüler-Interaktionen [Indirect messages about ability estimations in hypothetical teacher-student interactions]. *Zeitschrift für Pädagogische Psychologie*, 2, 235-242.
- Meyer, W.-U., Mittag, W., & Engler, U. (1986). Some effects of praise and blame on perceived ability and affect. *Social Cognition*, 4, 293-308. doi:10.1521/soco.1986.4.3.293
- Meyer, W.-U., & Plöger, F. O. (1979). Scheinbar paradoxe Wirkungen von Lob und Tadel auf die wahrgenommene eigene Begabung [Apparently paradoxical effects of praise and blame on the perceived talent]. In S. H. Phillip (Ed.), *Selbstkonzept-Forschung: Probleme, Befunde und Perspektiven* (pp. 221-236). Stuttgart: Klett-Cotta.
- Meyer, W.-U., Reizenzein, R., & Dickhäuser, O. (2004). Inferring ability from blame: Effects of effort- versus liking-oriented cognitive schemata. *Psychology Science*, 46, 281-293.
- Miller, A. T., & Hom, H. L., Jr. (1996). Conceptions of ability and the interpretation of praise, blame, and material rewards. *Journal of Experimental Education*, 65, 163-177. doi:10.1080/00220973.1997.9943790
- Möller, J. (1999). Soziale, fachbezogene und temporale Vergleichsprozesse bei der Beurteilung schulischer Leistungen: „Paradoxe“ Wirkungen von Lob und Tadel [Social, dimensional, and temporal comparisons: „Paradoxical“ effects of praise and blame]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31, 11-17. doi:10.1026//0049-8637.31.1.11
- Reizenzein, R. (1990). Gegen eine Fehlinterpretation des attributionstheoretischen Modells scheinbar paradoxer Effekte von Lob und Tadel [Against a misinterpretation of the attributional model of the seemingly paradoxical effects of praise and blame]. *Sprache & Kognition*, 9, 130-142.
- Reizenzein, R., & Battmann, W. (1992). Zu einigen (scheinbaren) methodischen Problemen bei der Überprüfung von Erklärungsmodellen der scheinbar paradoxen Wirkungen von Lob und Tadel: Antworten auf Groeben und Blickle [Notes on some (seemingly) methodological problems in the validation of models explaining seemingly paradoxical effects of

praise and blame: Responses to Groeben and Blickle]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39, 689-692.

Reisenzein, R., Debler, W., & Siemer, M. (1992). Der Verstehensvorgang bei scheinbar paradoxen Wirkungen von Lob und Tadel [The interpretation process in the seemingly paradoxical effects of praise and criticism]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39, 129-150.

Rheinberg, F., & Weich, K.-W. (1988). Wie gefährlich ist Lob? Eine Untersuchung zum „paradoxen Effekt“ von Lehrersanktionen [How dangerous is praise? A study about the paradoxical effect of teacher sanctions]. *Zeitschrift für Pädagogische Psychologie*, 2, 227-233.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods and the logic of conversation*. Hillsdale, NJ: Erlbaum.

Tacke, G., & Linder, F. (1981). Der Einfluß individualisierenden Lehrerverhaltens auf das Selbstkonzept von Schülern [The impact of individual teacher behaviour on the self concept of students]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 13, 190-193.

Weiner, B., & Kukla, A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, 15, 1-20. doi:10.1037/h0029211

## About the Authors

**Kai Kaspar** is a research associate and lecturer at the University of Osnabrück, Germany. He is member of the Institute of Psychology and of the Institute of Cognitive Science. In his research he focuses on bottom-up and top-down processes in human attention and perception, on mechanisms of embodied cognition, as well as on their relations to performance in everyday life and in human–computer interaction settings. The paradoxical effect of praise and blame – scrutinized in the present paper – is part of his research in the field of social interactions.

**Helena Stelz** is an occupational psychologist at the Education Authority of Lower Saxony, Germany. Her thematic priority is the teachers' health. She advises and supports state employees in primary, secondary, and vocational schools in recognizing health promotion of teachers as an important area of quality assurance in schools. Topics are team development and cooperation, conflict resolution, supervision, coaching and peer counseling, stress-, time- and selfmanagement, as well as healthfull leadership. The aim of occupational psychological counseling is to make the working conditions in school so that typical psychological stressors for the teaching profession are reduced.