Neural patterns underlying social comparisons of personal performance

Michael Lindner,[1,2,*] **Sarah Rudorf,**[1,*] **Robert Birg,**[1] **Armin Falk,**[1] **Bernd Weber,**[1,3,4] **and Klaus Fliessbach**[1,5,6]

[1]Center for Economics and Neuroscience, University of Bonn, Nachtigallenweg 86, 53127 Bonn, Germany, [2]School of Psychology and Clinical Language Sciences, University of Reading, Reading RG6 7BE, United Kingdom, [3]Department of Epileptology, University Hospital Bonn, Sigmund-Freud-Str. 25, 53105 Bonn, Germany, [4]Life&Brain Center, Department of NeuroCognition, University of Bonn, Sigmund-Freud-Str. 25, 53105 Bonn, Germany, [5]Department of Psychiatry, University Hospital Bonn, Sigmund-Freud-Str. 25, 53105 Bonn, Germany and [6]German Center for Neurodegenerative Diseases (DZNE), Sigmund-Freud-Str. 25, 53105 Bonn, Germany

Humans often evaluate their abilities by comparing their personal performance with that of others. For this process, it is critical whether the comparison turns out in one's favor or against it. Here, we investigate how social comparisons of performance are encoded and integrated on the neural level. We collected functional magnetic resonance images while subjects answered questions in a knowledge quiz that was related to their profession. After each question, subjects received a feedback about their personal performance, followed by a feedback about the performance of a reference group who had been quizzed beforehand. Based on the subjects' personal performance, we divided trials in downward and upward comparisons. We found that upward comparisons correlated with activity in the dorsolateral prefrontal cortex and the anterior insula. Downward comparisons were associated with increased activation in the ventral striatum (VS), the medial orbitofrontal cortex and the ventral anterior cingulate cortex (ACC). The extent to which subjects outperformed the reference group modulated the activity in the VS and in the dorsal ACC. We suggest that the co-activation of the VS and the dorsal ACC contributes to the integration of downward comparisons into the evaluation of personal performance.

Keywords: downward comparison; fMRI; self-evaluation; social comparison; performance

In line with social comparison theory, humans tend to compare themselves with others. The subject of these comparisons, i.e. the comparison dimension, varies widely across, e.g. abilities, opinions, personality traits, health or income. Presumably, the cognitive processes and the underlying neural mechanisms behind comparisons in these different dimensions also vary substantially. In this paper, we focus on abilities as a particularly prominent example for social comparisons (cp. first axiom in Festinger, 1954: "There exists, in the human organism, a drive to evaluate his opinions and his abilities."). Abilities are manifested through performance, and often times the only objective judgment of a performance is to compare it with other persons' performances. As such, performance comparisons provide the basis for university evaluation systems, sports competitions or the scoring of psychological performance tests to name but a few examples.

Recently, a line of research has investigated neural correlates of social comparison processes using functional magnetic resonance imaging (fMRI). Most of these studies have used monetary payoffs as the comparison dimension. These studies have consistently demonstrated the sensitivity of reward system activity in the ventral striatum (VS) and/or medial orbitofrontal cortex (mOFC) to relative income differences (Fliessbach et al. 2007; Bault et al. 2011; Dvash et al. 2010; Vostroknutov et al. 2012; Kang et al. 2013). This indicates that downward-directed social comparisons are encoded similarly to positive outcomes in reward learning. In most of the studies dealing with monetary reward differences, these differences were arbitrarily manipulated by the experimenter and did not depend on performance. Thus, these studies offer limited information of how personal performance is integrated in the social comparison process. One important exception is a study by Vostroknutov et al. (2012), which specifically showed that

mOFC activation depended on monetary reward inequalities more when they were related to performance than when they were related to luck. This study suggests that the mOFC plays an important role in integrating information about the magnitude of a social reward difference and its origin (whether it is due to luck or ability/performance). Our goal in the present study was to identify brain regions that encode social comparisons when objective performance differences provide the main comparison dimension.

In our experiment, we asked medical students to participate in a multiple-choice quiz on medical knowledge that was required for their intermediate examinations. After each question, subjects first received a feedback about their own performance and were then given the information of how many members of a reference group had answered this question correctly. We motivated our subjects to engage in social comparisons by using a homogenous and closely related reference group of other medical students, based on the assumption that individuals prefer to compare themselves with similar others (Suls et al. 2002). When a subject answered a question correctly, the advantage relative to the reference group performance was inversely proportional to the number of group members who had answered the question correctly (downward comparison deviation). Conversely, in the case of an incorrect answer, the disadvantage was proportional to the number of group members that had answered correctly (upward comparison deviation).

Our main research questions were whether downward and upward social comparisons of performance evoke distinct neural activation and whether this activation is proportional to the deviation between the personal performance and the reference group performance. We hypothesized that favorable downward comparisons of performance are encoded similarly to positive outcomes in reward-learning brain areas like the VS and that the degree of favorableness modulates this brain activation. Analogously, we assumed that unfavorable upward comparisons cause negative emotions, and we expected increased activity in brain regions that have been implicated in the processing of negative emotions, such as the anterior insula. Because social comparison theory assumes that the evaluation of personal abilities subserves a

general need for objective self-evaluation, we expected the involvement of regions that have previously been implied in self-evaluation (e.g. cortical midline structures) in the processing of both downward and upward comparisons.

MATERIAL AND METHODS
Participants

A total of 51 medical students from the Medical School of the University of Bonn participated in the study. All subjects had successfully passed the intermediate medical examination ('Physikum'), were right-handed (except one of the reference group) and gave written informed consent prior to the study. The study was approved by the ethics committee of the University of Bonn.

Twenty of these subjects (8 female; mean 23.85 ± 2.21 s.d. years) performed a computerized quiz with 150 multiple choice questions on basic medical knowledge and served as the reference group. The remaining 31 subjects took part in the fMRI experiment in which they performed a subset of 100 questions from the same quiz. One of these subjects had to be excluded because of technical problems, so that 30 subjects were finally analyzed (12 female, mean 23.93 ± 5.16 s.d. years). The two groups did not differ in years of study (mean 7.95 ± 2.39 s.d. vs mean 7.9 ± 2.54 s.d. years).

Stimuli and procedure
Reference group

Subjects in the reference group participated in a computerized quiz with 150 multiple-choice questions about concrete detailed knowledge of different medical fields that were all subject of the intermediate medical examination. Subjects were asked to choose the correct answer from a list of four alternative answers by pressing a button (for timing and an example question, see Figure 1). After each answer subjects received accuracy feedback before the next trial started. These subjects constituted the reference group for the fMRI experiment. They received an expense allowance of €10 and an additional €0.10 for each correct answer.

fMRI experiment

A different group of subjects was invited to the fMRI experiment. Prior to scanning, participants were shown photographs of the reference group members in order to increase awareness and credibility of the fact that the reference group consisted of real and comparable students. For the fMRI experiment, we selected 100 of the 150 questions from the reference group quiz. The selection of questions aimed at achieving a relatively even distribution of mean reference group accuracy across the full range of observed values. For this purpose, we selected the questions in the following way: first, we categorized the questions concerning their accuracy across subjects (10 categories). Then we performed the following step 50 times: from the category with the highest number of questions, we randomly chose one to exclude. The percentage of correct answers in the reference group for the remaining 100 questions ranged from 40 to 100 (median 65, quartiles 52.5, 80).

Subjects were confronted with these 100 questions in the fMRI experiment. The first part of each trial was identical to the procedure in the reference group quiz: after subjects had answered a question, they received an accuracy feedback about their own performance. For the fMRI group, the feedback about the personal performance was followed by a feedback about the reference group performance: The percentage of reference group members that had answered this question correctly/incorrectly was shown both graphically and numerically (see Figure 1). Subjects in the fMRI group received an expense allowance of €15 and an additional €0.10 for each correct answer.

With the feedback about the reference group performance, we operationalized social downward and upward comparisons in our task: when subjects performed correctly (incorrectly) then the feedback about the reference group performance yielded a downward (upward) comparison of performance. Further, the number of reference group members who had answered correctly defined the degree of deviation between subjects' personal performance and the reference group performance. For example, if a subject was correct (personal performance 100%) and received the information that 80% of the reference group members were correct, the deviation between the personal and the group performance would be $1 - 0.80 = 0.20$, yielding a downward comparison that is more favorable; the fewer reference group members were correct. In contrast, if a subject was incorrect (0%) and learned that 80% of the reference group members were correct, the deviation would be $0 - 0.80 = -0.80$, yielding an upward comparison that is less favorable; the more reference group members were correct.

fMRI data acquisition

Scanning was performed on a 1.5 Tesla Avanto Scanner (Siemens, Erlangen, Germany) using a standard 8 channel head coil. Slices were in axial orientation and covered all of the brain including the midbrain but not the entire cerebellum. For the functional scans, we collected echo planar images (EPI) with the following parameters: slice thickness 3 mm; inter slice gap 0.3 mm; matrix size 64×64; voxel size $3\,mm \times 3\,mm \times 3\,mm$; field of view $192\,mm \times 192\,mm$; echo time 50 ms; repetition time 2.5 s.

fMRI data preprocessing and analysis

FMRI data analysis was performed using the MATLAB (Mathworks)-based software Statistical Parametric Mapping 8 (SPM8, www.fil.ion. ucl.ac.uk/spm/). For preprocessing, functional images were realigned to the first image of each time series and again realigned to the mean image after first realignment. Images were then slice-time corrected using a sinc interpolation, normalized to the canonical EPI template used in SPM8 and smoothed with an 8 mm Gaussian kernel.

We modeled the blood oxygen level dependent signal (BOLD) response with a general linear model (GLM) that was estimated using a hemodynamic response function and a high-pass filter of 128 Hz as well as correction for autocorrelations. For this GLM, we defined regressors for five events (see Supplementary Figure S1): display of the question, display of the four choice options, time of the button press, feedback about the personal performance and feedback about the reference group performance. For all events except the display of the question, regressors were split into trials in which a subject's answer was correct and trials in which a subject's answer was incorrect. At the onset of the choice options, task difficulty as measured by the percentage of incorrect answers in the reference group and reaction times were added as parametric modulators. Task difficulty was also added as a parametric modulator at the onset of the personal performance feedback (see Supplementary Figure S5 for a *t*-test). This is crucial because we assumed that task difficulty influences the BOLD response at this time point and might therefore confound the effects of social comparison at the feedback about the reference group performance in the next screen: for example, in a difficult trial, an incorrect answer might be attributed to the difficulty, and the trial might be interpreted as less self-relevant; in contrast, a correct answer might be more self-esteem enhancing (Blackwood *et al.* 2003; Seidel *et al.* 2010) in a difficult than in an easy trial. We orthogonalized the regressors to each other in ascending order in our GLM to rule out that difficulty confounds the effect of the following main regressors of interest.

Our main regressors of interest modeled the two social comparison types when the feedback about the reference group performance was
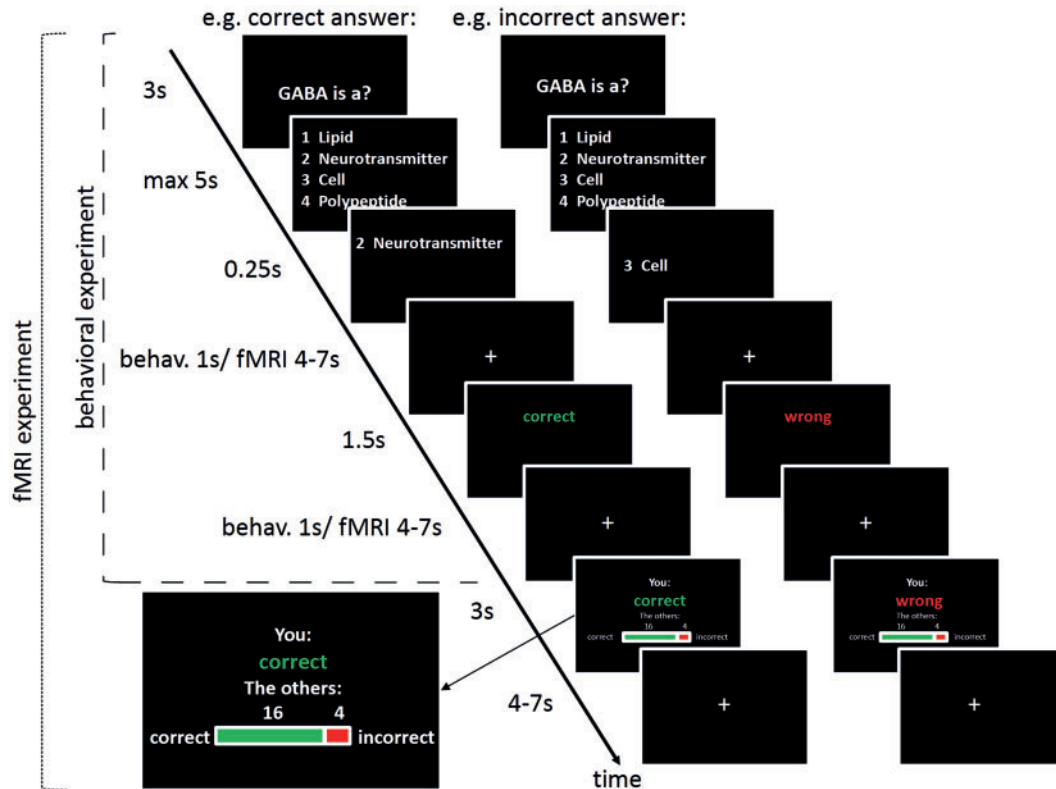
Fig. 1 Task design of the knowledge quiz for the reference group and the fMRI experiment. Each trial begins with the presentation of the question (3 s), followed by the four-answer alternative (max 5 s, terminated by response). After the subject has made a choice, the chosen answer is displayed for 250 ms. After a jittered interstimulus interval (ISI), the subject sees an accuracy feedback about the personal performance (1.5 s), followed by another ISI. The task for the reference group ends here. In the fMRI experiment, the subject subsequently sees a feedback about the reference group performance (3 s), followed by another jittered ISI.

displayed. We defined trials in which the subject's answer was correct as downward comparisons (performing better than the reference group), and trials in which the subject's answer was incorrect as upward comparisons (performing worse than the reference group). Note that the feedback about subjects' personal performance was given prior to the feedback about the reference group performance, and that regressors were orthogonalized in ascending order so that the variance explained by both events should be independent. We assume, however, that social comparison represents an integration of the personal and the reference group feedback. Therefore, we expected an overlap of activated regions for the difference between the social comparison types and the difference between correct and incorrect personal performance feedback (see Supplementary Figure S4). For each social comparison type, we added a parametric modulator (M_{down}, M_{up}) for the deviation Δ between the personal performance (correct, 1; incorrect, 0) and the percentage of correct answers in the reference group P_{ref}:

For performance deviations in downward comparisons,

$$M_{down} = \Delta(1 - P_{ref}),$$

and for performance deviations in upward comparisons,

$$M_{up} = \Delta(0 - P_{ref}).$$

In our sample, this resulted in parameter values of M 0.3 ± 0.17 s.d., range 0–0.6, for the modulator M_{down}, and values of M -0.59 ± 0.14 s.d., range -1–-0.4, for the modulator M_{up} (see Supplementary Figure S3 for an illustration of the parameter distributions for all individuals). To identify brain regions that are differentially activated by the two types of comparison, we computed a paired t-test between the parameter estimates for the onsets of downward and

upward comparisons. To account for the possibility that the relative number of downward comparison trials per subject (see Supplementary Figure S2 for a histogram) confounds the contrast between downward and upward comparisons, we included the relative number of downward comparison trials as a covariate in the paired t-test at the group level. To further test which brain region's activity is modulated by the deviation between personal performance and reference group performance, we ran two separate one-sample t-tests on the respective parametric modulators (M_{down}, M_{up}) at the display of the reference group performance feedback for downward and upward comparisons. Based on our a priori hypothesis that downward comparisons are encoded in reward system structures, we performed small-volume corrections for the ventral midbrain, the VS and the mOFC, using masks derived from the Harvard–Oxford atlas (Frazier *et al.* 2005; Desikan *et al.* 2006; Makris *et al.* 2006; Goldstein *et al.* 2007). Because the anatomical structures concerning our other hypotheses (e.g. brain structures involved in the processing of self-referential information) are less well characterized, we refrained from applying further region-specific analyses and conducted whole-brain analyses instead. All results are corrected for multiple comparisons using family-wise error (FWE) correction at a statistical threshold of $P < 0.05$, either small-volume corrected for reward system structures or on the whole-brain level. We visualized the results using the SPM toolboxes xjView (www.alivelearn.net/xjview) and rfxplot (Gläscher 2009).

RESULTS

Behavioral results

The 20 participants of the reference group showed, across all 150 questions, a mean accuracy of $54.83 \pm 0.12\%$ s.d. and a mean reaction time

of 3.44 ± 0.41 s.d. For the selected set of 100 questions for the fMRI experiment, the mean accuracy was $65.75 \pm 0.11\%$ s.d., and the mean reaction time was 3.34 ± 0.39 s s.d. A *t*-test between the reaction times of the correct (mean 3.13 ± 0.4 s s.d.) and incorrect trials (mean 3.81 ± 0.41 s s.d.) showed that subjects took longer to respond in incorrect than correct trials ($T_{(19)} = 11.463$, $P < 0.01$). The 30 participants in the fMRI experiment answered correctly in a mean of $63.83 \pm 9.5\%$ s.d. of all trials with a mean reaction time of 3.75 ± 0.37 s s.d. A *t*-test between the reaction times of the correct and incorrect trials revealed a significant difference ($T_{(29)} = 16.158$, $P < 0.01$) with greater reaction times for incorrect (mean 4.36 ± 0.42 s s.d.) than for correct trials (mean 3.42 ± 0.32 s s.d.). Mann–Whitney–Wilcoxon tests between the reference group ($n_1 = 20$) and the fMRI group ($n_2 = 30$) revealed a significant group difference for the overall reaction time (U $= 132$, $z = 3.24$, $P < 0.01$ two-tailed), but not for accuracy (U $= 322.5$, $z = 0.436$, $P > 0.05$ two-tailed).

fMRI results

Downward and upward comparisons of performance involve distinct neural structures

We first tested whether downward and upward comparisons of performance correlated with BOLD activity in different brain regions, while controlling for each subject's relative number of downward comparison trials. A paired *t*-test showed that downward comparisons yielded stronger activation than upward comparison in the ventral anterior cingulate cortex (ACC) (whole-brain corrected $P_{FWE} < 0.05$; Figure 2, in red), the mOFC and the bilateral VS (small-volume corrected $P_{FWE} < 0.05$). The ventral midbrain showed no significant effect. The opposite contrast of upward greater than downward comparisons demonstrated strong bilateral activation of the dorsolateral prefrontal cortex (PFC), the anterior insula and the dorsomedial PFC (whole-brain corrected $P_{FWE} < 0.05$; Figure 2, in blue). Table 1 reports the peak voxel coordinates of these activations. To test whether the contrast downward *vs* upward comparisons at the time of the reference group feedback shows a similar response pattern to the contrast correct *vs* incorrect personal performance feedback, we ran a paired *t*-test of the respective contrast images. As expected, there is an overlap of activation in the VS and the ACC, and deactivation in the dorsolateral PFC (see Supplementary Figure S4). Taken together, these results show that distinct brain regions integrate information about personal and others' performance during downward and upward comparisons.

Neural responses to downward comparisons are proportional to favorableness

Next, we were interested in whether the BOLD activation for social comparisons is proportional to the deviation between the personal performance and the reference group performance. We used this deviation as a parametric regressor at the onset of social upward and downward comparisons, respectively. For the negative deviation from the reference group performance at upward comparisons, no significant modulation of the BOLD signal was observed. For downward comparisons, on the other hand, we found that the BOLD signal in the bilateral VS (small-volume corrected $P_{FWE} < 0.05$) and in the dACC (whole-brain corrected $P_{FWE} < 0.05$) increased significantly with the favorableness of the comparison, i.e. the positive deviation from the reference group performance (Figure 3; Table 2).

DISCUSSION

In this study, we provide first evidence for distinct neural signals of social downward and upward comparisons during performance evaluation. Critically, we used a performance task that allowed us to provide

subjects with an objective measure of their personal performance and to compute the deviation of personal performance from a reference group's performance. We found that downward comparisons led to increased activity in brain regions commonly associated with reward processing. Moreover, activity in the VS and the dACC correlated with positive deviation of personal performance from the reference group performance. Upward comparisons, on the other hand, were associated with increased activity in the anterior insula, the dorsolateral and dorsomedial PFC.

Our results suggest that downward-directed social comparisons of performance are encoded similarly to positive outcomes in reward processing. When we contrasted the two comparison types (regardless of the deviation from the reference group) while controlling for inter-individual differences in the relative number of downward comparison trials, we found that the VS and mOFC, brain areas related to reward learning (Knutson *et al.* 2001; O'Doherty 2004; Schultz and Tremblay 2006), encode downward comparisons. Conversely, upward comparisons evoked responses in the dorsolateral PFC, the anterior insula and the medial PFC, areas that have been associated with cognitive control and the processing of motivationally important events (Craig 2009; Hare *et al.* 2009; Singer *et al.* 2009). This is principally consistent with the notion that an upward comparison is an emotionally salient event that can motivate a change in goal-directed behavior. Taken together, these results reveal candidate regions for the distinct processing of upward and downward social comparisons of performance.

The VS, and especially its component the nucleus accumbens, is a key structure in the dopaminergic reward system. It encodes information in a reference-point dependent manner, i.e. it signals discrepancies between outcomes and expectations. In classical learning theory, reference points result from the learning history, but recent studies have given many examples of how reference points can also be provided by social information such as the income (Fliessbach *et al.* 2007) or reputation of others (Meshi *et al.* 2013). Our study complements this previous research by showing that social downward comparisons of performance, as a non-material comparison dimension, evoke reference-dependent processes in the VS that are similar to the processing of primary and secondary rewards.

The dACC is part of the cortical midline structures that have been consistently attributed to the processing of self-related information (for a meta-analysis, see Northoff *et al.* 2006). Previous research has reported activation in the dACC when subjects evaluate their own personality traits, but also when they evaluate the traits of others (Ochsner *et al.* 2005; Qin and Northoff 2011). Interestingly, activity in this region is modulated by the perceived dissimilarity between personal and others' opinions or preferences (Mitchell *et al.* 2006; Tamir and Mitchell 2010). This has been interpreted to reflect a judgment process in which individuals use their own opinion as a reference to infer other people's opinions. Other studies report a modulation of dACC activity by uncertainty in non-social domains, i.e. the expected difference between outcome and expectation (Krain *et al.* 2006; Rudorf *et al.* 2012). In our task, the reference group's error rate could serve as an indicator for uncertainty of the performance evaluation. This would imply a more general role of the dACC during the comparison with others. It has also been suggested that increasing dissimilarity between the self and others increases the uncertainty of self-evaluation, and that dACC activity reflects this uncertainty (Flagan and Beer 2013). In our study, increased dACC activity during more favorable downward comparisons might therefore encode the salience of this information and signal a higher degree of certainty about own abilities. Assume a subject in our experiment solves a task correctly. The feedback that all comparison group members solved the task correctly would contain little information about the own ability, and having solved the task correctly would appear as a matter of course. Whereas the feedback
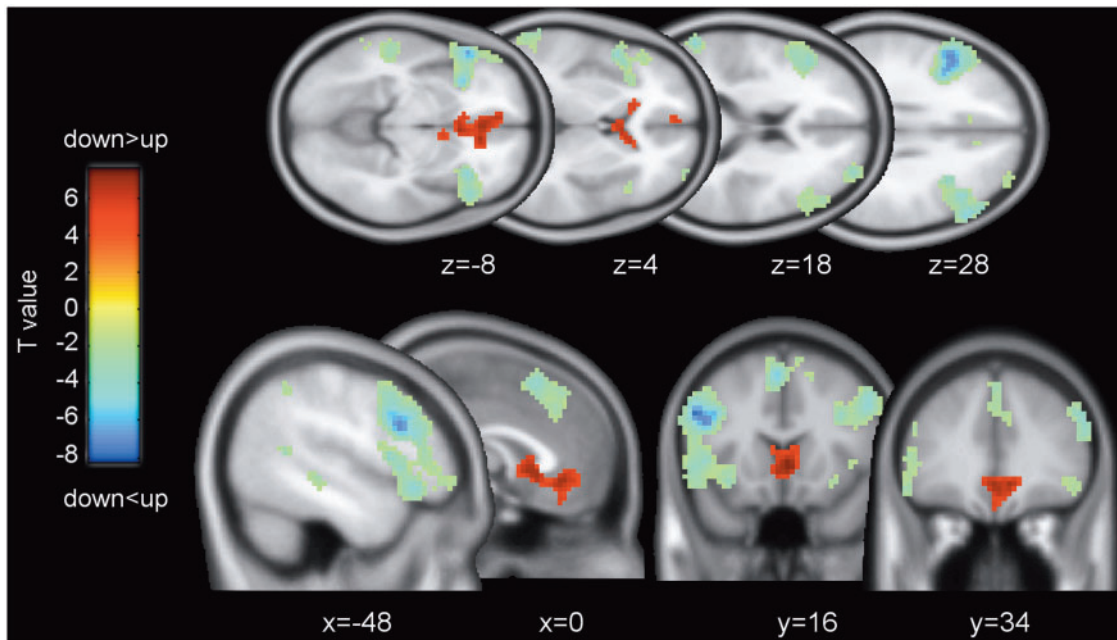
Fig. 2 Main effect of social comparison direction. Neural activations of the contrast of downward greater than upward comparisons are shown as T-maps ($P < 0.001$, uncorrected, $k > 10$, for demonstration purposes). The effect of 'downward > upward' is displayed as positive T-values (red); the opposite effect of 'downward < upward' is displayed as negative T-values (blue). T-values are color-coded as specified by the color bar.

Table 1 Main effect of social downward and upward comparisons

Type of comparison	Region	Laterality	MNI coordinates			Cluster size k_E	Maximum *t*-value
			x	*y*	*z*		
Downward *vs* upward comparison	Ventral ACC	R	3	20	−5	334	6.00**
	Medial OFC	R	9	35	−8		5.25***
		L	(−6)	(50)	(−11)		(4.13*)
			(0)	(38)	(−14)		(3.94*)
	VS	R	6	8	−5		5.17***
			(9)	(17)	(−2)		(3.69*)
		L	(−6)	(14)	(−5)		(3.91*)
Upward *vs* downward comparison	IFG/ aIns/ dlPFC	L	−54	17	31	1174	9.83**
			−42	14	31		9.21**
			−54	26	−8		8.30**
	dlPFC	R	57	29	31	526	7.37**
			48	23	31		6.51**
			54	38	22		5.83***
	dmPFC	L	−3	17	55	300	7.13**
			0	35	49		4.59***
			0	32	40		4.56***
	aIns/IFG	R	36	26	−8	219	6.48**
			51	26	−5		5.87***
			30	14	−14		4.08***
	Superior temporal gyrus	L	−60	−55	19	179	6.30**
	Inferior parietal lobe	L	−36	−52	46	270	5.49***
	dlPFC	R	33	59	19	131	5.49***
	Medial temporal gyrus	L	−54	−31	−8	116	4.10***

Results from the random effects analyses are shown. Deviating peak voxels from the small-volume correction are shown in parentheses. Height threshold, $t = 3.40$; extent threshold, 100 voxels. L, left; R, right; ACC, anterior cingulate cortex; OFC, orbitofrontal cortex; IFG, inferior frontal gyrus; aIns, anterior insula; dlPFC, dorsolateral prefrontal cortex; dmPFC, dorsomedial prefrontal cortex.
*$P < 0.05$. The activation survives small-volume correction for multiple comparisons based on FWE control at the peak level for a priori-defined regions of interest (masks derived from the Harvard–Oxford atlas). Inclusion threshold, $t = 3.40$.
**$P < 0.05$. The activation survives whole-brain correction for multiple comparisons based on FWE control at the peak level.
***$P < 0.05$. The activation survives whole-brain correction for multiple comparisons based on FWE control at the cluster level.
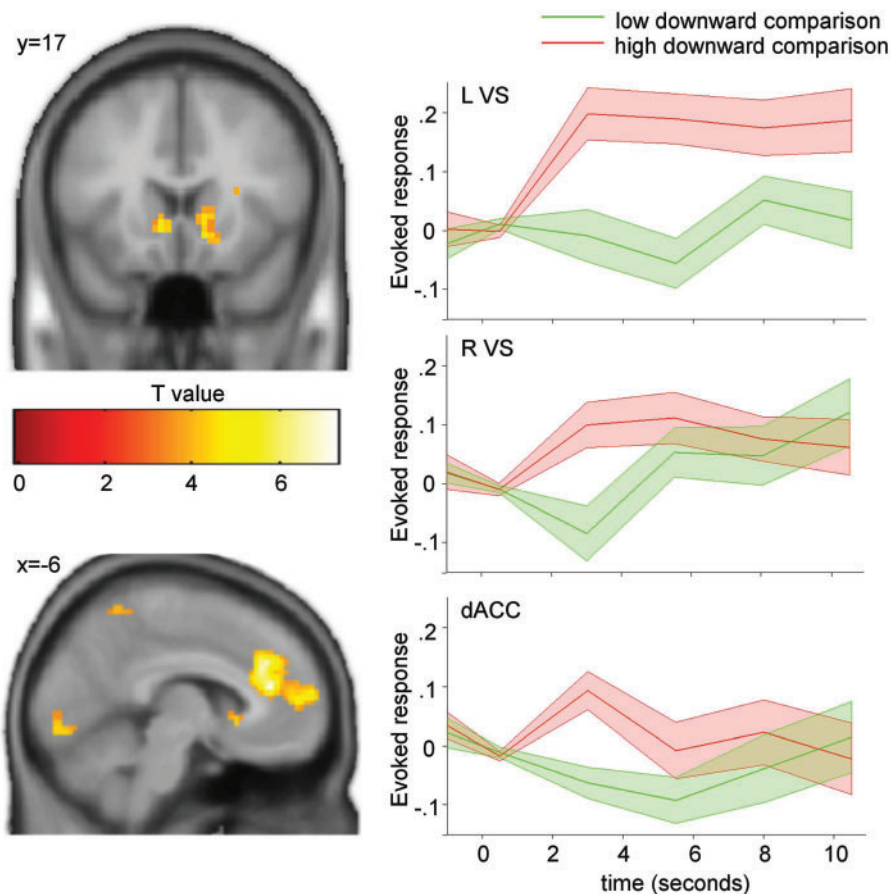
Fig. 3 Deviation between personal and reference group performance in downward comparisons. Neural activations in the bilateral VS and the dACC increase with the positive deviation of the personal performance from the reference group performance during downward comparisons. Left, T-maps for downward comparisons modulated by the degree of deviation (P < 0.001, unc., k > 10, for demonstration purposes). T-values are color-coded as specified by the color bar. Right, peri-stimulus time histograms in the respective regions split into lower and upper percentiles of downward comparisons. Colored areas show standard errors of the mean. L, left; R, right; VS, ventral striatum; dACC, dorsal anterior cingulate cortex.

Table 2 Effect of deviation degree (downward comparison)

Region	Laterality	MNI coordinates			Cluster size k_E	Maximum t-value
		x	y	z		
dACC	L/R	−6	38	16	440	7.34**
Occipital cortex	L/R	−12	−97	−5	266	6.59**
VS	R	15	8	−11	216	5.90*
		(12)	(8)	(−11)		(5.74*)
	L	−9	23	−2	15	5.61*
		(−9)	(17)	(−5)		(5.01*)
dlPFC	L	−27	35	34	55	4.88***
	R	15	50	31	17	4.35[†]
Superior temporal gyrus	R	72	−25	25	53	4.88***
	L	−54	11	−5	47	4.51***
Precuneus	R	3	−58	52	71	4.55***

Results from the random effects analysis are shown. Deviating peak voxels from the small-volume correction for the VS are shown in parentheses. Height threshold, t = 3.40; extent threshold, 15 voxels. L, left; R, right; dACC, dorsal anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex.
*P < 0.05. The activation survives small-volume correction for multiple comparisons based on FWE control at the peak level for a priori-defined regions of interest (masks derived from the Harvard–Oxford atlas). Inclusion threshold, t = 3.40.
**P < 0.05. The activation survives whole-brain correction for multiple comparisons based on FWE control at the peak level.
***P < 0.05. The activation survives whole-brain correction for multiple comparisons based on FWE control at the cluster level.
[†]P < 0.001. The activation does not survive correction for multiple comparisons. For the sake of completeness the activation is reported at P < 0.001, uncorrected.

that no comparison group members solved the task correctly would be more salient and implicate that the good own performance can be attributed to high own ability with greater certainty. The involvement of other dACC functions such as error detection and performance monitoring is also possible, but probably less dominant because there was no trial-by-trial learning in our task. In line with the reviewed literature, it seems plausible that subjects use the information about others' performance to evaluate their own abilities. Given its general involvement in uncertainty coding and its sensitivity to performance deviation in downward comparisons in our task, one role of the dACC during social comparisons might be to promote the acquisition of a stable representation of personal abilities.

A previous study by Vostroknutov et al. (2012) also addressed the impact of performance (vs luck) on social comparison processes. In contrast to our study, the comparison dimension was not performance itself but different monetary payoffs based on skill or effort (which are the principle factors determining performance). The authors showed a specific involvement of the mOFC when payoff differences occurred in a skill condition (vs a luck condition), suggesting that this brain region integrates information on outcomes and the way this outcome was achieved. This is in line with a recent finding that the modulation of the BOLD signal in the mOFC by reward magnitude strongly depends on the history of a reward also in a non-social condition (Hernandez-Lallement et al. 2014). In contrast to these findings, the involvement of the dACC in our study might be due to the fact that the main focus of

comparison was on the performance itself instead of the resulting outcome. However, as performance was also linked to a monetary reward in our study (see caveats below), future studies are needed to confirm this.

In a previous study on cultural differences in social comparisons of monetary rewards, Kang et al. (2013) observed the involvement of a more anterior part of cortical midline structures (the ventromedial PFC) only in Korean but not in American subjects. Thus, the sensitivity of cortical midline structures to social differences might also critically depend on cultural differences.

Finally, a differential co-activation of reward-related and self-related brain regions for upward and downward comparisons has not been observed in the majority of previous neuroimaging studies investigating social comparisons. Critically, our experimental design differs from these previous studies in two ways: the self-relevance of the comparison dimension and the objectivity of the reference point. Some studies implemented arbitrary monetary rewards as comparison dimension and did not show an involvement of cortical midline structures (Fliessbach et al. 2007; Bault et al. 2011), potentially because arbitrary rewards have a low impact on self-evaluation. In other studies where subjects evaluated their own personality traits (Hughes and Beer 2013) and future expectations (Blair et al. 2013), the comparison dimension was self-relevant, but subjects did not have an objective reference point they could relate the reference group information to. Moreover, personality traits and future expectations are complex and ambiguous constructs whose self-evaluation is biased by self-enhancement and overoptimism, and this complexity might explain why those studies found activation of a larger set of brain regions associated with emotion processing and regulation such as the insula, the amygdala and the lateral PFC. Our study, in contrast, used a simple, objective measure of personal performance (performance in a medical knowledge quiz) that is also highly relevant for self-evaluation (abilities as a medical student). Our data suggest that when the result of a social comparison is in one's favor and this result is integrated to inform the self-evaluation of personal performance, a concordant activation of VS and dACC takes place. This interpretation is in line with a study by Korn et al. (2012). They observed a positive concurrent modulation of activity in the VS and the dACC/medial PFC by the degree of favorableness of other subjects' ratings of personality traits. They did not find a modulation by the deviation of others' ratings from the personal ratings. It should be noted, however, that their analysis collapsed positive and negative deviations. In our analysis, we tested positive and negative deviations separately and found that activity in the VS and dACC was modulated by positive, but not negative deviations from the reference group. Together with previous findings our results suggest that the processing of self-enhancing information about a personal characteristic in the light of external social information is mediated by a co-activation of VS and dACC.

For upward comparisons of performance, we observed no such reference-point dependent modulation. Of course, a null finding renders any interpretation speculative. One speculation is that the diverging results between downward and upward comparisons reflect a fundamental bias in humans toward a positive self-evaluation. This self-serving bias might result in a preference for positive information provided by social downward comparisons and a neglect of negative information provided by upward comparisons. A possible approach to further investigate this issue would be to run similar experiments in individuals known to have a reduced self-serving bias such as patients with depression (Seidel et al. 2012). A methodological reason for this null finding might also be that our sample includes less upward than downward comparison trials (M 36.17 vs M 63.83). This is due to the fact that events depended on real answers. In our sample, the minimum number of upward comparison trials in a single subject was 18,

and only 9 subjects experienced <30 upward comparisons. Moreover, the range of parameter values for the upward performance deviations was not any narrower than that of downward deviations. However, it is possible that the parametric analysis of upward comparisons lacked the statistical power to detect a significant BOLD effect. It would be interesting to see whether a study that manipulates the number of comparison events could find an effect for upward comparisons.

This study was a first approach to investigate neural processes underlying social comparisons of objective performance. We would like to acknowledge the caveats of this study and give suggestions on how to address open questions in the future. First of all, our task was designed so that upward comparisons occurred only after incorrect personal performance and downward comparisons only after correct performance. Although the feedback about the reference group performance was temporally separated from the feedback about the personal performance, we cannot fully exclude that the processes during the feedback about others' performances still partially reflect reminiscence about the own performance. Future studies are therefore needed to replicate these results while controlling for personal performance, e.g. by keeping performance constant and adding information about the performance of others that is either superior (upward comparison) or inferior (downward comparison). We addressed the potential confound of reminiscence about personal performance at the time of social comparison statistically and performed a parametric analysis. Here, we tested separately for upward and for downward comparisons whether BOLD activity was modulated by the deviation between personal performance and reference group performance. For upward comparisons, no such modulation was observed, whereas for downward comparisons, we found a positive modulation of BOLD activity in the VS and the dACC by favorableness of the comparison. In other words, when subjects were informed about the reference group performance after they had solved the task correctly themselves, activity in these regions was greater; the fewer reference group members had also solved the task. One possible explanation is that this pattern of brain activation reflects two interacting components of the social comparison process: reward and self-evaluation. This hypothesis could be tested in future studies by including behavioral measures of perceived reward and self-evaluation.

Second, the fact that we did not counterbalance the colors for correct vs incorrect performance during the reference group feedback might, theoretically, contribute to the modulation in the BOLD signal. This seems unlikely, however, given the fact that there is no known (directed) color dependency of brain responses in our regions of interest.

Third, we did not manipulate self-relevance in our task, e.g. by including a control task with a less self-relevant dimension. So we can only assume that our performance measure (medical quiz) and the comparison process were self-relevant to our subjects (medical students).

Finally, although our study aimed at performance as the main comparison dimension, we did use small monetary incentives to motivate subjects to show optimal performance. Thus, it is possible that at the time of the social feedback, subjects compared their "income" with the average group income. In contrast to previous studies, however, payoffs were kept constant with performance, so we assume that our results are mainly driven by performance differences.

Despite these caveats, this study gives useful insights into a sparsely studied subdomain of social comparisons, and it provides a number of concrete hypotheses that can be tested in future research.

In summary, our study highlights distinct candidate brain regions for social upward and downward comparisons of performance. Our results suggest a co-activation of the VS and dACC as a potential neurophysiological basis for the integration of positive information

about personal performance into the representation of own abilities. This neural pattern therefore provides a potential marker for the investigation of maladaptive social comparison processes, which are well known to occur, e.g. in depression or eating disorders (Green *et al.* 2009).

SUPPLEMENTARY DATA

Supplementary Data are available at *SCAN* online.

Conflict of Interest

None declared.

REFERENCES

Bault, N., Joffily, M., Rustichini, A., Coricelli, G. (2011). Medial prefrontal cortex and striatum mediate the influence of social comparison on the decision process. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 16044–9.

Blackwood, N.J., Bentall, R.P., ffytche, D.H., Simmons, A., Murray, R.M., Howard, R.J. (2003). Self-responsibility and the self-serving bias: an fMRI investigation of causal attributions. *NeuroImage*, *20*(2), 1076–1085.

Blair, K.S., Otero, M., Teng, C., et al. (2013). Dissociable roles of ventromedial prefrontal cortex (vmPFC) and rostral anterior cingulate cortex (rACC) in value representation and optimistic bias. *NeuroImage*, *78*, 103–10.

Craig, A.D.B. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*, 59–70.

Desikan, R.S., Ségonne, F., Fischl, B., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*, 968–80.

Dvash, J., Gilam, G., Ben-Ze'ev, A., Hendler, T., Shamay-Tsoory, S.G. (2010). The envious brain: the neural basis of social comparison. *Human Brain Mapping*, *31*(11), 1741–50.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–40.

Flagan, T., Beer, J.S. (2013). Three ways in which midline regions contribute to self-evaluation. *Frontiers in Human Neuroscience*, *7*, 450.

Fliessbach, K., Weber, B., Trautner, P., et al. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, *318*, 1305–8.

Frazier, J.A., Chiu, S., Breeze, J.L., et al. (2005). Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. *The American Journal of Psychiatry*, *162*, 1256–65.

Gläscher, J. (2009). Visualization of group inference data in functional neuroimaging. *Neuroinformatics*, *7*, 73–82.

Goldstein, J.M., Seidman, L.J., Makris, N., et al. (2007). Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability. *Biological Psychiatry*, *61*, 935–45.

Green, M.A., Scott, N.A., Cross, S.E., et al. (2009). Eating disorder behaviors and depression: a minimal relationship beyond social comparison, self-esteem, and body dissatisfaction. *Journal of Clinical Psychology*, *165*, 989–999.

Hare, T., Camerer, C.F., Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, *324*, 646–8.

Hernandez-Lallement, J., Kuss, K., Trautner, P., Weber, B., Falk, A., Fliessbach, K. (2014). Effort increases sensitivity to reward and loss magnitude in the human brain. *Social Cognitive and Affective Neuroscience*, *9*, 342–9.

Hughes, B.L., Beer, J.S. (2013). Protecting the self: The effect of social-evaluative threat on neural representations of self. *Journal of Cognitive Neuroscience*, *25*, 613–22.

Kang, P., Lee, Y., Choi, I., Kim, H. (2013). Neural evidence for individual and cultural variability in the social comparison effect. *The Journal of Neuroscience*, *33*(41), 16200–8.

Knutson, B., Fong, C.A.G.W., Adams, C.M., Varner, J.L., Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*, 3683–7.

Korn, C.W., Prehn, K., Park, S.Q., Walter, H., Heekeren, H.R. (2012). Positively biased processing of self-relevant social feedback. *The Journal of Neuroscience*, *32*, 16832–44.

Krain, A.L., Wilson, A.M., Arbuckle, R., Castellanos, F.X., Milham, M.P. (2006). Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. *NeuroImage*, *32*, 477–84.

Makris, N., Goldstein, J.M., Kennedy, D., et al. (2006). Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophrenia Research*, *83*, 155–71.

Meshi, D., Morawetz, C., Heekeren, H.R. (2013). Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. *Frontiers in Human Neuroscience*, *7*, 439.

Mitchell, J.P., Macrae, C.N., Banaji, M.R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*, 655–63.

Northoff, G., Heinzel, A., De Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J. (2006). Self-referential processing in our brain–a meta-analysis of imaging studies on the self. *NeuroImage*, *31*, 440–57.

Ochsner, K.N., Beer, J.S., Robertson, E.R., et al. (2005). The neural correlates of direct and reflected self-knowledge. *NeuroImage*, *28*, 797–814.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology*, *14*, 769–76.

Qin, P., Northoff, G. (2011). How is our self related to midline regions and the default-mode network? *NeuroImage*, *57*, 1221–33.

Rudorf, S., Preuschoff, K., Weber, B. (2012). Neural correlates of anticipation risk reflect risk preferences. *The Journal of Neuroscience*, *32*, 16683–92.

Schultz, W., Tremblay, L. (2006). Involvement of primate orbitofrontal neurons in reward, uncertainty, and learning. In: Zald, D.H., Rauch, S.L., editors. *The Orbitofrontal Cortex*. Oxford: Oxford University Press, pp. 173–98.

Seidel, E.-M., Eickhoff, S.B., Kellermann, T., Schneider, F., Gur, R.C., Habel, U., Derntl, B. (2010). Who is to blame? Neural correlates of causal attribution in social situations. *Social Neuroscience*, *5*, 335–50.

Seidel, E.-M., Satterthwaite, T.D., Eickhoff, S.B., Schneider, F., Gur, R.C., Wolf, D.H., Habel, U., Derntl, B. (2012). Neural correlates of depressive realism — An fMRI study on causal attribution in depression. *Journal of Affective Disorders*, *138*, 268–76.

Singer, T., Critchley, H.D., Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, *13*, , 334–40.

Suls, J., Martin, R., Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, *11*, 159–63.

Tamir, D.I., Mitchell, J.P. (2010). Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 10827–32.

Vostroknutov, A., Tobler, P.N., Rustichini, A. (2012). Causes of social reward differences encoded in human brain. *Journal of Neurophysiology*, *107*(5), 1403–12.