Evidence for impaired attention to social stimuli in autism has been mixed. The role of social feedback in shaping attention to other, non-social stimuli that are predictive of such feedback has not been examined in the context of autism. In the present study, participants searched for a color-defined target during a training phase, with the color of the target predicting the emotional reaction of a face that appeared after each trial. Then, participants performed visual search for a shape-defined target while trying to ignore the color of stimuli. On a subset of trials, one of the non-targets was rendered in the color of a former target from training. Autistic traits were measured for each participant using the Autism Quotient (AQ). Our findings replicate robust attentional capture by stimuli learned to predict valenced social feedback. There was no evidence that autistic traits are associated with blunted attention to predictors of social outcomes. Consistent with an emerging body of literature, our findings cast doubt on strong versions of the claim that autistic traits can be explained by a blunted influence of social information on the attention system. We extend these findings to non-social stimuli that predict socially relevant information.

Keywords: selective attention; social feedback; reward learning; autism quotient

Only a small fraction of the information our perceptual systems are exposed to influences our behavior (e.g., Mack & Rock, 1998; Rensink, O’Regan, & Clark, 1997). Limitations in the brain’s capacity to represent perceptual input create conditions under which stimuli compete for representation at later, capacity-limited stages of information processing (Desimone & Duncan, 1995). Attention serves as the mechanism by which the brain selects which among multiple stimuli receive such representation (Desimone & Duncan, 1995).

In order for an individual to function in society, attention must select stimuli that provide information that is useful for guiding behavior (e.g., Corbetta & Shulman, 2002). One important source of such information is the reactions of other people. Autistic individuals struggle to use this information effectively (e.g., Dawson et al., 2004; Kanner, 1943; Warlaumont et al., 2014), which raises important questions concerning the relationship between the traits that are characteristic of autism and how the attention system is influenced by social information.

An influential hypothesis concerning the etiology of autism centers on the role of attention to social stimuli and the motivational properties of social feedback. Specifically, it is hypothesized that social feedback fails to evoke a reward response that is sufficient to facilitate social behavior in autism, which results in a tendency to ignore social information (Chevallier et al., 2012; Dawson et al., 1998, 2005; Schultz, 2005). This reduction in the time spent attending to social information then has a cascading impact on social learning, resulting in the impoverished development of basic processes such as face and speech perception (Grelotti, Gauthier, & Schultz, 2002; Kuhl et al., 2005; Schultz et al., 2000).

Evidence for impaired attention to social stimuli in autism has been decidedly mixed. While some studies have provided evidence for reduced attention to social stimuli (Chawarska, & Shic, 2009; Chawarska, Volkmar, & Klin, 2010; Chevallier et al., 2012; Dawson et al., 1998, 2005; Kikuchi et al., 2011; Schultz, 2005), cases of unimpaired attention, with intact and robust preferences for social stimuli, have also been reported (Elsabbagh et al., 2013; Fischer et al., 2014; Fletcher-Watson et al., 2008; New et al., 2010; Sheth et al., 2011; van der Geest et al., 2001). It seems not to be the case that autism can be explained simply by a broad tendency to ignore social information across situations and contexts. Here, we were interested in whether a different but related aspect of social attention might provide a more sensitive indicator of the traits characteristic of autism.

One of the ways in which we learn the value and importance of objects in our environment is through positive or negative social feedback (e.g., Izuma & Adolphs, 2013; Goldstein & Schwade, 2008; Palmer & Schloss,
Sensitivity to such stimulus-outcome contingencies would be important for guiding pro-social and other adaptive behaviors. Should social feedback fail to shape attentional priorities in an individual, broad deficits in social reciprocity might be expected, as well as blunted preferences for socially-relevant stimuli. Aberrant feedback processing is a well-documented feature of autism, particularly when outcomes are never fully predictable or rely on complex contextual contingencies (e.g., Dawson et al., 2001; Larson et al., 2011; Van de Cruys et al., 2014; Vlaming et al., 2008). The role of social feedback in shaping attention to other, non-social stimuli that are predictive of such feedback has not been examined in the context of autism.

Autistic traits vary across individuals and can be measured at sub-clinical levels in the normal population (Baron-Cohen et al., 2001). A well-validated measure for quantifying the severity of autistic traits is the Autism Quotient (AQ; Baron-Cohen et al., 2001). AQ scores have been shown to predict both attention measures (gaze cuing: Bayliss, di Pellegrino, & Tipper, 2005; Bayliss & Tipper, 2005; global information processing: Grinter et al., 2009) and brain structure and functioning (Nummenmaa et al., 2012; von dem Hagen, 2011), in a manner consistent with differences observed between autistic patients and controls. In the present study, we quantified the severity of autistic traits within the sub-clinical range using the AQ, and related this measure to the magnitude of attentional biases towards stimuli associated with valenced social outcomes. We predicted that autistic traits would be inversely correlated with attentional biases towards stimuli associated with valenced social outcomes, due to difficulty learning from social outcomes, broader difficulties in learning to predict probabilistic outcomes (e.g., Dawson et al., 2001; Larson et al., 2011; Van de Cruys et al., 2014; Vlaming et al., 2008), and/or a failure to prioritize stimuli based on social learning.

The role of associative learning in the control of attention has been well-documented. Stimuli that have been learned to predict a monetary or food reward automatically draw attention (Anderson et al., 2011a, 2011b, 2014a; Anderson & Yantis, 2013; Pool et al., 2014; Yantis et al., 2012; see Anderson, 2016b, for a recent review), as do stimuli associated with aversive outcomes such as electric shock (Schmidt, Belopolsky, & Theeuwes, 2015a, 2015b; Wang, Yu, & Zhou, 2013). Importantly, similar biases can result from valenced social feedback: Arbitrary (non-social) stimuli that are consistently paired with positive (Anderson, 2016a) or negative (Anderson, 2017) social feedback have been shown to automatically draw attention in healthy young adults.

In the present study, participants first completed a training phase in which two different target colors were associated with either a high or a low probability of valenced social feedback, thus serving as a predictive cue for such feedback. In a subsequent test phase, the degree to which these colors automatically capture the attention of participants was examined. We replicate attentional biases for stimuli that were reliably followed by either positive (happy) or negative (angry) facial expressions during training, and relate the magnitude of these measured biases to the severity of autistic traits as measured by the AQ.

**Methods**

**Participants**

181 participants were recruited from the Texas A&M University community, 84 (25 male, 58 female [1 not reported], mean age = 18.8 y) in the happy emotion condition and 97 (16 male, 76 female [5 not reported], mean age = 19.1 y) in the angry emotion condition. All reported normal or corrected-to-normal visual acuity and normal color vision. Participants were compensated with course credit. Data from an additional 8 participants were discarded due to withdrawal from the study before completing the experimental task, failure to complete the entire AQ, or chance-level performance (accuracy < 60%). Data collection for each condition was stopped at the end of the week that 80 participants was reached, a number which would yield the power to detect correlations as small as $r = \pm 0.22$. All participants provided written informed consent, and all study procedures were approved by the Texas A&M University Institutional Review Board and conformed to the principles outlined in the Declaration of Helsinki.

**Apparatus**

A Dell OptiPlex equipped with Matlab software and Psychophysics Toolbox extensions (Brainard, 1997) was used to present the stimuli on a Dell P2717H monitor. The participants viewed the monitor from a distance of approximately 70 cm in a dimly lit room. Manual responses were entered using a standard keyboard.

**Autism Quotient**

Each participant completed the Autism Quotient (AQ; Baron-Cohen et al., 2001) survey prior to completing the experimental task. Responses were scored in terms of the number of autism-consistent statements endorsed by the participant, with scores ranging from 0–50. Although the questionnaire is not appropriate for diagnostic purposes, a score of 32 or greater is predictive of clinical autism (Baron-Cohen et al., 2001).

**Training Phase**

**Stimuli.** Each trial consisted of a fixation display, a search array, and a feedback display (Figure 1A). The fixation display contained a white fixation cross (0.5° × 0.5° visual angle) presented in the center of the screen against a black background, and the search array consisted of the fixation cross surrounded by six colored circles (each 2.3° × 2.3°) placed at equal intervals on an imaginary circle with a radius of 5°. The target was defined as the red or green circle, exactly one of which was presented on each trial (red on 50% of trials and green on the remaining 50%); the color of each nontarget circle was drawn from the set {blue, cyan, pink, orange, yellow, white} without replacement. Inside the target circle, a white bar was oriented either vertically or horizontally, and inside each of the nontargets, a white bar was tilted at 45° to the left or to the right (randomly determined for each nontarget). The feedback display consisted of a
picture of a face exhibiting either a valenced or a neutral expression. For the happy expression condition, the faces were those of 20 male and 20 female models taken from the AR face database (Martinez & Benavente, 1998). For the angry expression condition, the faces were those of 12 male and 12 female models taken from the Warsaw Set of Emotional Facial Expression Pictures (Olszanowski et al., 2015). Both face sets were photographs of real people modelling a variety of expressions, of which happy and neutral and angry and neutral, respectively, were used for the feedback display in the present study.

**Design.** One of the two color targets (alternating across participants) was followed by a face exhibiting a valenced expression on 80% of trials and a face exhibiting a neutral expression on the remaining 20% (high-valence target); for the other color target, these percentages were reversed (low-valence target). In both the happy and angry expression conditions, the same models were used for the valent and neutral faces (i.e., each model had a valent and neutral counterpart), such that the gender and identity of the faces that communicated the social feedback were balanced across the two target conditions. Each individual face appeared equally-often. Each color target appeared in each of the six possible stimulus locations equally-often, and trials were presented in a random order.

**Procedure.** The training phase consisted of 240 trials, which were preceded by 40 practice trials. Each trial began with the presentation of the fixation display for a randomly varying interval of 400, 500, or 600 ms. The search array then appeared and remained on screen until a response was made or 1000 ms had elapsed, after which the trial timed out. The search array was followed by a blank screen for 1000 ms, the feedback display for 1500 ms, and a blank 1000 ms inter-trial interval (ITI).

Participants were instructed to find the circle that was either red or green on that trial, and to identify the orientation of the bar within this red or green target. To identify the bar within the target, participants were instructed to press the “z” key if the bar was oriented vertically and the “m” key if the bar was oriented horizontally. They were instructed to respond both quickly and accurately. The nature of the feedback following each search array was independent of the participants’ actual behavior; that is, it was not affected by the speed or accuracy of the response on that (or any) trial. Participants were only informed that the faces would “react to what happened on each trial.” If the trial timed out, the words “Too Slow” were centrally presented for 1000 ms.

**Test Phase**

**Stimuli.** Each trial consisted of a fixation display, a search array, and (in the event of an incorrect response) a feedback display (Figure 1B). The six shapes now consisted of either a diamond among circles or a circle among diamonds, and the target was defined as the unique shape. On a subset of the trials, one of the nontargets was rendered in the color of a former target from the training phase (referred to as the distractor); the target was never red or green. The feedback display only informed participants if their prior response was incorrect.

**Design.** Target identity, target location, distractor identity, and distractor location were fully crossed and counterbalanced, and trials were presented in a random order. Distractors were presented on 50% of the trials, half of which were high-valence distractors and half of which were low-valence distractors (high- and low-valence color from the training phase, respectively).

**Procedure.** Participants were instructed to ignore the color of the shapes and to focus on identifying the unique shape both quickly and accurately, using the same orientation-to-response mapping. The test phase consisted of 240 trials, which were preceded by 32 practice...
(distractor-absent) trials. In the event of an incorrect response, the search array was followed immediately by the word “Incorrect” centrally presented for 1000 ms; no faces were shown during the test phase. Each trial ended with a 500 ms ITI. Trials timed out after 1500 ms. As in the training phase, if the trial timed out, the words “Too Slow” were centrally presented for 1000 ms.

Data Analysis
Only correct responses were included in all analyses of RT, and RTs more than 3 SDs above or below the mean of their respective condition for each participant were trimmed. Analyses of behavioral data focus on RT, which this paradigm is most sensitive to (e.g., Anderson et al., 2011b, 2013, 2014b; Anderson, 2016a, 2017); we had no specific predictions concerning accuracy. The magnitude of attentional capture by social cues, or valence-driven attentional capture, was defined as the difference in RT between the high-valence distractor and distractor-absent conditions, as has been frequently used in studies examining individual differences in learning-dependent attentional capture (e.g., Anderson et al., 2011b, 2013, 2014b, 2016a, 2016b; Anderson & Yantis, 2012; Qi et al., 2013). We related this measure of attentional capture to autistic traits as measured by the AQ. Follow-up analyses examining the correlation between valence-driven attentional capture and AQ score separately for each training condition utilize Bonferroni correction for two comparisons ($\alpha = 0.025$). To the degree that autistic traits are associated with a blunting of social attention, a negative correlation would be predicted.

Results

**Autism Quotient**
The mean AQ for participants in the happy (mean = 17.2, stdev = 5.21, range = 4–30) and angry (mean = 18.0, stdev = 6.19, range = 7–35) expression conditions were comparable, $t(179) = 0.95$, $p = 0.344$. The mean and variability of this measure were similar to those reported in a meta-analysis of AQ scores in the normal population (Ruzich et al., 2015).

**Training Phase**
RT data were submitted to a 2 × 2 analysis of variance (ANOVA) with target valence (high vs low) as a within-participants factor and the type of emotion (happy vs angry) as a between-participants factor. This analysis revealed no reliable effects, main effect of valence: $F(1,179) = 2.35$, $p = 0.127$, main effect of emotion: $F(1,179) = 0.83$, $p = 0.362$, interaction: $F(1,179) = 0.62$, $p = 0.433$. The same ANOVA performed on accuracy data yielded similar results, main effect of valence: $F(1,179) = 2.03$, $p = 0.156$, main effect of emotion: $F(1,179) = 0.11$, $p = 0.738$, interaction: $F(1,179) = 0.88$, $p = 0.350$ (see Table 1).

**Test Phase**
RT data were submitted to a 3 × 2 ANOVA with distractor condition (absent, low-valence, high-valence) as a within-participants factor and the type of emotion (happy vs angry) as a between-participants factor. This analysis revealed a highly robust effect of distractor condition, $F(2,358) = 9.05$, $p < 0.001$, $\eta^2_p = 0.048$ (see Figure 2A). Planned pairwise comparisons revealed that the high-

![Figure 2](https://online.ucpress.edu/colabra/article-pdf/4/1/10/468177/119-1846-3-pb.pdf)

**Table 1**: Mean response time and accuracy by target condition during the training phase. Standard errors of the mean are in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Happy</th>
<th>Angry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-valence</td>
<td>High-valence</td>
</tr>
<tr>
<td>Response Time (ms)</td>
<td>620 (5.0)</td>
<td>616 (5.4)</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>85.1 (1.3)</td>
<td>85.3 (1.1)</td>
</tr>
</tbody>
</table>
valence distractors slowed responses compared to both the low-valence distractor, \(t(180) = 3.26, p = 0.001, d = 0.25\), and distractor-absent conditions, \(t(180) = 4.11, p < 0.001, d = 0.31\). The difference in RT between the low-valence distractor condition and the distractor-absent condition was not significant, \(t(180) = 0.04, p = 0.968\).

The main effect of emotion, \(F(1,179) = 0.09, p = 0.768\), and the interaction, \(F(2,358) = 0.51, p = 0.604\), were not significant. The same ANOVA performed on accuracy data did not reveal any significant effects, main effect of valence: \(F(2,358) = 0.13, p = 0.878\), main effect of emotion: \(F(1,179) = 1.71, p = 0.193\), interaction: \(F(2,358) = 1.13, p = 0.324\) (see Table 2).

Across both emotion conditions, there was no evidence for a negative correlation between valence-driven attentional capture (difference in RT between high-valence distractor and distractor-absent conditions) and autistic traits, \(r = 0.052, p = 0.483\). However, separately examining participants by training condition (happy vs angry) revealed that the magnitude of the correlation differed significantly between the two training conditions, \(z = 2.34, p = 0.019\), suggesting that collapsing across this variable was not warranted. Considering each training condition separately, participants in the happy expression condition actually showed a small but reliable positive correlation, \(r = 0.258, p = 0.018\) (significant with Bonferroni correction), while participants in the angry expression condition did not show a significant correlation, \(r = -0.091, p = 0.373\) (Figure 2B). A similar pattern of results was obtained using the difference between the high-valence and low-valence distractor conditions as the measure of attentional capture, \(r = 0.231, p = 0.034\) (marginally significant with Bonferroni correction) vs \(r = -0.133, p = 0.195\), direct comparison: \(z = 2.43, p = 0.015\). Autistic traits were unrelated to the difference in accuracy between the high-valence distractor condition and either of the two other distractor conditions, \(rs < 0.086, ps > 0.25\).

**Discussion**

In the present study, we first replicate evidence that stimuli associated with both positive (Anderson, 2016a) and negative (Anderson, 2017) social feedback automatically capture attention. Consistent with many prior studies, no effects of feedback were evident during the training phase (e.g., Anderson, 2016a, 2017; Anderson et al., 2011a, 2013, 2014b, 2016a; Anderson & Halpern, 2017), suggesting that performance was dominated by the influence of task goals in this part of the task. However, when goals and attention to valenced stimuli conflict, as in the test phase, a robust attentional bias was evident.

The present study also offers a direct comparison between attention to positively- and negatively-valenced social cues. With large sample sizes, there was no evidence for a negativity bias or a reward bias (the JZS Bayes factor for the comparison is 5.63 in favor of the null hypothesis, which is considered very strong evidence; see Rouder et al., 2009). Neither social reinforcement nor punishment appears to be generally more privileged in its ability to modulate attention. An interesting question for future research remains whether these two sources of attentional bias reflect a common underlying mechanism, or whether they reflect independent mechanisms with a similar behavioral profile.

Rather than measure attention to social stimuli, such as faces, which has been more extensively studied in the context of autism and autistic traits (e.g., Chawarska, & Shic, 2009; Chawarska, Volkmar, & Klin, 2010; Chevallier et al., 2012; Dawson et al., 1998, 2005; Elsabbagh et al., 2013; Fischer et al., 2014; Fletcher-Watson et al., 2008; Kikuchi et al., 2011; New et al., 2010; Schultz, 2005; Sheth et al., 2011; van der Geest et al., 2001), our experimental task was specifically designed to measure the ability of social feedback to shape attention to the stimuli that predict such feedback. Specifically, we measured the degree to which attention is biased towards stimuli that are associated with a high probability of being followed by valent reactions. An inability to adjust attentional priorities due to learning from social feedback, rather than attend to social stimuli per se, may offer a more sensitive indicator of autistic traits, which is a possibility that we set out to test here.

We found little evidence for the idea that autistic traits are associated with blunted attentional biases towards stimuli that predict socially valenced outcomes. A significant negative correlation would have been expected if autistic traits are associated with (a) difficulty learning relationships between social outcomes and neutral stimuli and/or (b) a weaker influence of such learning on the attention system. Our findings are inconsistent with either of these predictions. If anything, there was a slight positive correlation between autistic traits and attentional biases towards positively-valenced social cues, although this is clearly not reflective of a robust and generalizable phenomenon as it was not replicated using negatively-valenced social cues. We therefore hesitate to draw any firm conclusions regarding this relationship, although we note that seemingly greater attention to socially-relevant stimuli in autism is not without precedent (Elsabbagh et al., 2013), and it does suggest intact learning from social feedback across the normal range of autistic traits.

**Table 2**: Mean response time and accuracy by distractor condition during the test phase. Standard errors of the mean are in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Happy</th>
<th></th>
<th></th>
<th>Angry</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Low-valence</td>
<td>High-valence</td>
<td>Absent</td>
<td>Low-valence</td>
<td>High-valence</td>
</tr>
<tr>
<td>Response Time (ms)</td>
<td>717 (9.5)</td>
<td>716 (9.2)</td>
<td>729 (9.0)</td>
<td>721 (8.1)</td>
<td>722 (8.2)</td>
<td>730 (7.9)</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>86.9 (0.8)</td>
<td>86.6 (0.9)</td>
<td>86.3 (0.9)</td>
<td>87.7 (0.7)</td>
<td>87.6 (0.7)</td>
<td>88.3 (0.7)</td>
</tr>
</tbody>
</table>
In the present study, autistic traits were measured within the range typical of a normal college-aged adult population. Scores within the sub-clinical range of this measure of autistic traits have been related to both patterns of attentional orienting (gaze cuing: Bayliss et al., 2005; Bayliss & Tipper, 2005; global information processing: Grinter et al., 2009) and brain structure and functioning (Nummenmaa et al., 2012; von dem Hagen, 2011). Furthermore, sub-clinical scores on a different scale, the Beck Depression Inventory (Beck, Steer, & Brown, 1996), have been linked to differences in attentional capture using a similar experimental paradigm (Anderson et al., 2014b, 2017), supporting the sensitivity of our attention measure. Care should be taken, however, in generalizing our findings beyond the range of autistic traits observed in the normal population. It is possible that clinically-significant autism might present with a qualitatively different attention profile than the range observed here, and recent evidence suggests that variability in autistic traits may have a categorical structure with a distinct sub-type reflecting clinically-significant impairment (James et al., 2016).

It is important to distinguish between attentional biases driven by associative learning and attentional biases driven by former target history (Anderson & Halpern, 2017; Sha & Jiang, 2016). In our experimental design, the critical distractors were not only previously associated social feedback during training, but these stimuli were also previously task-relevant. Therefore, a tendency to attend to these stimuli could be explained by a bias to select former targets independently of any feedback-related processing. The difference in RT between the high- and low-valence conditions, however, argues specifically in favor of an associative learning account. Distractors in both of these conditions were former targets, which were searched for equally-often, but they differed in their probability of being followed by valent social feedback. We therefore conclude that attention was biased predominantly by associative learning between color stimuli and social feedback in our task, and that such biases were not negatively correlated with autistic traits.

As stated previously, our experimental task focused specifically on the influence of the differential valence of social feedback in shaping attentional biases. A different, but related, question concerns attention to stimuli that do and do not predict social interaction. Perhaps individuals high in autistic traits show a reduced preference for stimuli that predict any face vs stimuli that predict either a non-social outcome or no outcome. To examine this possibility, future studies could employ a similar paradigm in which one color target is consistently followed by a neutral face during training while another color target is never followed by a social stimulus.

Consistent with an emerging body of literature (e.g., Fischer et al., 2014; Elsabbagh et al., 2013; Fletcher-Watson et al., 2008; New et al., 2010; Sheth et al., 2011; van der Geest et al., 2001), our findings cast doubt on strong versions of the claim that autistic traits can be explained by an attention deficit related to socially-relevant information. With a large sample size, our findings lend further credence to these negative results, and extend them to non-social stimuli that predict socially-relevant information.

### Data Accessibility Statement
All data from the reported experiment are available as supplemental material linked to this manuscript.

### Additional Files
The additional file for this article can be found as follows:

- **Spreadsheet S1.** Raw data from the experiment (XLSX file). DOI: https://doi.org/10.1525/collabra.119.s1

### Acknowledgements
The authors are grateful to A. Montgomery, H. Begum, S. Weisssgerber, and R. Vance for assistance with data collection.

### Competing Interests
The authors have no competing interests to declare.

### Author Contributions
BAA and HK contributed to the conception and design of the study, HK organized data collection, BAA and HK analyzed and interpreted the data, BAA drafted the manuscript, BAA and HK revised the manuscript and approved the final version.

### References


Autism spectrum traits in the typical population predict structure and function in the posterior superior temporal sulcus. *Cerebral Cortex, 21*, 493–500. DOI: https://doi.org/10.1093/cercor/bhq062

