Preface and Acknowledgements

I would like to thank my supervisor, Robert Biegler, for introducing me to this topic. He has also introduced me to new theories, perspectives and methods, as well as created software for this project. Without his expertise there would simply be no thesis. Lastly, thank you introducing me to preregistrations, and for converting me to Bayesian statistics.

I would also like extend my gratitude to my co-supervisor Gerit Pfuhl, who has helped with the instrument development, data collection and analyses, as well as given me valuable feedback. I would also like to thank her students, as well as those of my friends, who have helped with recruitment and pilot testing. In addition, my gratitude goes to Carina Lisøy Glerud, for giving me helpful tips on style and format.

Lastly, I would like to thank Silje Eggen for her support throughout these two years. Our long discussions and brainstorming sessions have been immensely helpful and valuable.
Abstract

The phenotypic behaviour associated with psychosis and autism spectrum disorder have been theorised to stem from diametrically opposite tendencies in mental state attribution, thereby placing the phenotypes on opposite ends of a mentalising dimension. By extending this mentalising dimension to agency attribution, this model supplements Bayesian theories that argue that false prediction errors cause of behaviour associated with psychosis and autism spectrum disorder. The aim of the present study was therefore to investigate whether tendencies towards psychosis and autism spectrum disorder relates to opposite tendencies in agency attribution. A chasing detection task was modified for the present study. Participants from the general population completed the visual search task, in which the goal was to identify chasing motion in artificial stimuli. In line with predictions, psychotic traits were associated with a bias to see chasing. In contrast, no conclusive evidence was found that indicated that autistic traits were related to a bias to not see chasing. In addition, an oddity was identified when examining the relationship between bias and ability to detect chasing. The implication of these findings and recommendations for further research were discussed.
### Content

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Psychosis and Autism</td>
<td>3</td>
</tr>
<tr>
<td>Autism Spectrum Disorder</td>
<td>3</td>
</tr>
<tr>
<td>Traits in the General Population</td>
<td>4</td>
</tr>
<tr>
<td>Similarities and Comorbidity</td>
<td>5</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>6</td>
</tr>
<tr>
<td>A Dimensional Model</td>
<td>7</td>
</tr>
<tr>
<td>Mentalising</td>
<td>10</td>
</tr>
<tr>
<td>Agency Attribution</td>
<td>11</td>
</tr>
<tr>
<td>Problems and Gaps</td>
<td>13</td>
</tr>
<tr>
<td>The present study</td>
<td>16</td>
</tr>
<tr>
<td>Method</td>
<td>18</td>
</tr>
<tr>
<td>Participants</td>
<td>22</td>
</tr>
<tr>
<td>Measures and Procedure</td>
<td>22</td>
</tr>
<tr>
<td>Statistical Analysis and Preliminary Analysis</td>
<td>23</td>
</tr>
<tr>
<td>Results</td>
<td>26</td>
</tr>
<tr>
<td>Discussion</td>
<td>28</td>
</tr>
<tr>
<td>The Relationship Between IDAQ And Bias c</td>
<td>31</td>
</tr>
<tr>
<td>IDAQ, Bias c and Traits: Mental state attribution and Traits</td>
<td>31</td>
</tr>
<tr>
<td>The Relationship Between Traits</td>
<td>32</td>
</tr>
<tr>
<td>The Relationship between D’ and Proportion of Correct Identification</td>
<td>37</td>
</tr>
<tr>
<td>The Relationship between Ability and Bias c, and IDAQ</td>
<td>37</td>
</tr>
<tr>
<td>ID, d’ and Traits: Ability and Traits</td>
<td>40</td>
</tr>
<tr>
<td>Limitations and Strengths</td>
<td>40</td>
</tr>
<tr>
<td>Implications and Future Research</td>
<td>41</td>
</tr>
<tr>
<td>Conclusion</td>
<td>42</td>
</tr>
<tr>
<td>References</td>
<td>46</td>
</tr>
<tr>
<td>Appendices</td>
<td>60</td>
</tr>
</tbody>
</table>
Seeing Minds: A Signal Detection Study of Agency Attribution in Autism and Psychosis

When the first edition of the Diagnostic and Statistical Manual for Mental Disorders was published in 1952, autism was described as the psychotic manifestation of childhood schizophrenia (American Psychiatric Association, 1952). In his landmark paper from 1943, Leo Kanner presents excerpts where the behaviour of an autistic girl is described using phrases like “ignored persons completely” and “expression of ideas completely lacking” (p. 232). While such descriptions are today associated with autism spectrum disorder (ASD), this girl is one of the many autistic children who prior to the 1980s were diagnosed with schizophrenia. ASD and psychotic disorders have since then been classified into separate diagnoses. However, this earlier tendency to merge psychotic and autistic behaviour may not have been entirely unfounded. Although the behavioural expressions of ASD and psychosis are quite distinct, individual accounts have identified the same mechanism as the potential cause of autistic and psychotic symptoms.

In 2005, C. Frith developed a theory with the aim to explain the core features of psychosis; delusions and hallucinations. Delusions are bizarre, false beliefs that are not easily understood by individuals with the same culture or background (Fletcher & Frith, 2009). Hallucinations are false perceptions that occur in the absence of external stimuli (American Psychiatric Association, 2013). C. Frith (2005) proposed that both delusions on control and hallucinations can be explained by a consistent failure to make accurate prediction about the world. This results in unexpected and seemingly strange experiences. These experiences can only be explained by bizarre and highly abstract ideas, and external agents are often identified as the causes due to excessive agency attribution. Fletcher and Frith (2009) later extended this to include all delusions. These delusions become part of one’s model of the world, which then influences perceptual processing, thereby creating hallucinations.

The symptoms associated with ASD has also been linked to prediction. Van de Cruys et al. (2014) recently proposed that people on the autistic spectrum treat all unexpected events as important, failing to consider that some errors occur due to the stochastic nature of the world. The authors argue that these large deviances between what you expect would happen and what actually happens demand attention, as they signify that something important has to be learned. The world may appear as if continuously changing in important ways, resulting in sensory overload. Thus, social interactions (which as complex and noisy) are avoided, and
repetitive behaviour and specialised interests are used to create predictable environments as a countermeasure to sensory overload.

If autistic and psychotic behaviour can both be explained by consistently perceiving one’s predictions to be markedly different from one’s experiences, why are the behavioural expression so inherently different? One answer is linked to the social features of psychosis and ASD. Both psychosis and ASD have been linked to problems with mentalising (see, e.g., Kimhi, 2014; Pickup, 2006). Mentalising, or theory of mind, refers to the ability to make mental state attributions (Van Overwalle & Baetens, 2009). We make inferences about other peoples’ intentions, beliefs, thoughts and emotions, and use this information to explain and predict behaviour. Abu-Akel, Wood, Hansen, and Apperly (2015) demonstrated that tendencies towards psychosis and ASD can be placed on opposite sides of a mentalising dimension. Their findings suggest that poor mentalising ability is caused by different tendencies in mental state attribution; autistic tendencies are related to reduced mental state attribution, while psychotic tendencies are related to excessive mental state attribution. This is referred to as hypomentalising and hypermentalising, respectively.

How does this hypo- and hypermentalising dimension fit with the theories explaining psychotic and autistic behaviour? It may be that differences in mental state attribution is not what causes differences in behaviour, but rather what these differences imply. It is possible that one’s tendency to attribute mental states corresponds with one’s tendency to attribute agency. Abu-Akel et al.’s (2015) dimension would then predict excessive agency attribution in psychosis, as proposed by C. Frith (2005). It would also predict reduced agency attribution in ASD, although this idea was not discussed by Van de Cruys et al. (2014).

While agency attribution has previously been investigated in the context of psychotic or autistic tendencies, none have used signal detection analysis to examine whether these tendencies relate to distinct biases within the same detection paradigm. Findings that support such diametrically opposite biases would provide a basis to combine the theories of C. Frith (2005), Fletcher and Frith (2009) and Van de Cruys et al. (2014) to a collective theory of psychosis and autism, in which behavioural outcomes are linked to opposite biases in agency attribution. If this is justified, this could have implication for the management of autistic and psychotic symptoms. Early implementation of methods for identification and management of atypical agency attribution, perhaps even prediction errors, might reduce the expressions of ASD and psychosis. Thus, the aim of the present study is to augment research on this topic, by investigating how agency attribution relates to tendencies towards ASD and psychosis.
Psychosis and Autism

**Psychosis.** Psychosis is known to many as a loss of touch with reality. Although there are core elements that repeat within each class of psychotic disorders, psychotic symptoms vary in frequency and severity. When the fifth and newest version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V; American Psychiatric Association, 2013) was published, a dimensional approach to psychotic disorders was introduced. In contrast to the previous categorical approach, schizophrenia spectrum disorders (SSD) and other psychotic disorders were now placed on a psychotic spectrum based on the number, severity and duration of psychotic symptoms. By using this dimensional approach, psychotic disorders are understood based on the spectrum of symptoms rather than by subtypes of the schizophrenia diagnosis (Bhati, 2013).

Investigations of the prevalence of psychotic disorders show the prevalence rate varying from 0.7% to approximately 3% in the population (Hilker et al., 2018; Perala et al., 2007; van Nierop et al., 2012; van Os, Linscott, Myin-Germeys, Delespaul, & Krabbendam, 2009). By analysing data from a national sample of twins, Hilker et al. (2018) obtained a heritability estimate for SSDs at 73%. Although psychosis often is thought of as the core features of SSDs, the psychotic spectrum also extends to disorders like bipolar and depressive disorders that share these psychotic features (Crespi & Badcock, 2008). Specifically, the positive psychotic symptoms listed in the DSM-V are 1) delusions, 2) hallucinations, 3) disorganized thinking (inferred from speech) and 4) grossly disorganized or abnormal motor behaviour. The criteria for the several SSDs, including schizophrenia, requires the presence of at least one of symptom 1), 2) or 3). This has the consequence of limiting the literature review for the present research question, as the theoretical scope only entails delusions and hallucinations.

There is no shortage of research studies using schizophrenia-based samples, but these samples may be too heterogenous due to varying occurrence of symptoms within diagnoses. A review from 2014 show that the prevalence of visual hallucination in schizophrenia samples vary from 4% to 65%, and from 25% to 86% for auditory hallucinations (Waters et al., 2014). In a study from 2012 where people diagnosed with schizophrenia or schizoaffective disorder self-reported symptom occurrence for one week, only 52% reported having at least one delusional experience (Ben-Zeev, Morris, Swendsen, & Granholm, 2012). In fact, research on genotype-phenotype relationships suggests that schizophrenia can be divided into eight genetically distinct subtypes, each with their own sets of symptoms (Arnedo et al., 2015). Caution should therefore be exercised when reviewing literature that fail to control for psychotic symptoms.
In addition, negative symptoms, such as a reduced emotional expression and anhedonia, might also influence research results. Gooding and Pflum (2011) demonstrated that individuals displaying signs of positive schizotypy (perceptual aberration and magical thinking) perform poorer on a mentalising task than those displaying signs of negative schizotypy (social anhedonia). Studies like these underlines the importance of controlling for positive and negative symptoms in clinical samples, and might help explain mixed findings in the literature.

**Autism Spectrum Disorder.** Autism spectrum disorder (ASD), like psychotic disorders, has also been characterised along a continuum. In the DSM-V, the diagnosis of ASD entails the entire spectrum. However, the severity is determined by the level of impairment in social communication as well as the severity of the restricted, repetitive behaviours (American Psychiatric Association, 2013). Unlike SSD, ASD does not have subgroups in the DSM-V. A categorical approach still exists in the tenth revision of the International Classification of Diseases (ICD-10), where the spectrum is divided into diagnoses like Autistic disorder, Rett’s syndrome and Asperger’s syndrome.

Restricted, repetitive behaviours and impairments in reciprocal social behaviour are the core features of ASD (American Psychiatric Association, 2013). These are referred to as the non-social and social symptoms of ASD, respectively. The social symptoms refer to deficits in social communication and social interaction. These are central to the ASD diagnosis, yet adult females diagnosed with ASD show fewer socio-communication deficits than adult males (Lai et al., 2011). A gender differences in non-social symptoms (i.e. restricted, repetitive behaviours, interest or activities) has also been found when measuring symptoms in children (Mandy et al., 2012). Females show fewer symptoms of restricted, repetitive behaviour than males.

In a Swedish twin study, ASD was found to have a very high heritability of 80% (Lichtenstein, Carlstrom, Rastam, Gillberg, & Anckarsater, 2010). Elsabagh et al. (2012) reviewed prevalence rates collected worldwide, and found a prevalence rate of 0.6% for children with ASD. The prevalence rate in the adult population have been found to be slightly higher, at approximately 1% (Brugha et al., 2011). The latter study is not as extensive as the former, and one should therefore be cautious when comparing these numbers. In addition, changes in diagnostic criteria and learned compensation strategies in the elderly population also makes it difficult to assess whether the prevalence rate decrease or increase with age (American Psychiatric Association, 2013; Brugha et al., 2011; Kaland, 2018). Nevertheless, ASD is considered a lifelong condition (Kaland, 2018).
**Traits in the General Population.** Twin studies have shown that both the autism and psychosis phenotypes are heritable (Constantino & Todd, 2003; Linney et al., 2003; Ronald et al., 2006), leading to autism- and psychosis-spectrum traits being distributed in the general population. These traits are attenuated manifestations of the symptoms of ASD and SSD (Happe, Ronald, & Plomin, 2006; Murray, Booth, McKenzie, Kuenssberg, & O’Donnell, 2014; van Os et al., 2009). By measuring phenotypic traits in non-clinical samples, studies show that these traits are distributed along continua, with multiple correlated dimensions (Hoekstra et al., 2011; Stefanis et al., 2002). There is also evidence that those with SSD and ASD diagnoses represent the extreme end of the continuous distributions of subclinical traits (Lundstrom et al., 2012; Zavos et al., 2014).

Both positive, negative and disorganised subclinical traits of the psychosis phenotype is distributed in the general population (Dominguez, Saka, Lieb, Wittchen, & van Os, 2010; Werbeloff et al., 2015). These traits are frequently reported in non-clinical samples (De Loore et al., 2011; Nuevo et al., 2012; Rössler et al., 2015), making the psychosis phenotype more common than psychotic disorders (Linscott & van Os, 2013; van Nierop et al., 2012; van Os et al., 2009). Dominguez et al. (2010) found that 17% of non-clinical participants reported lifetime cumulative incidence of only positive psychotic traits, while 15.7% reported lifetime cumulative incidence of only negative/disorganised traits. Although common, these traits are often benign (Loewy, Johnson, & Cannon, 2007; Preti et al., 2014) and transitory (Hanssen, Bak, Bijl, Vollebergh, & van Os, 2005; Linscott & van Os, 2013).

However, there is evidence that negative and disorganised traits predict positive traits, and that these positive traits can become persistent and distressing when combined with negative traits (Dominguez et al., 2010). Persistent positive traits alone (Werbeloff et al., 2012), or in combination with negative traits (Werbeloff et al., 2015), have been linked to increased risk of clinical outcome later in life. The distress associated with a psychotic or psychotic-like experience is what predominantly determines the severity of that experience (Hanssen et al., 2003). van Os et al. (2009) make a distinction between subclinical psychotic experiences and subclinical psychotic symptoms, where only symptoms are associated with distress or help-seeking behaviour. In their systematic review, they found the rate of psychotic experiences to be higher (around 8%) than the rate of symptoms (around 4%) in the general population. Taken together with the even lower prevalence rate of SSD further down the psychosis continuum, this is evidence that the phenotype “splinters out” in the general population.
Population-based twin studies show that autistic traits are highly heritable (Ronald & Hoekstra, 2011). In a systematic review, Ruzich et al. (2015) found a normal distribution of autistic traits in the general population. This review also revealed that people diagnosed with ASD were found on the higher end of the continuum, which support the notion of a continuous distribution of traits extending from clinical to subclinical (but see also, Frazier et al., 2010). There is evidence that both social (Robinson et al., 2016) and non-social (Morrison, Chambers, Faso, & Sasson, 2018) features extends to the healthy end of the continuum. Ronald et al. (2006) found a modest correlation between traits reflecting social impairments, communication impairments and repetitive behaviours and interest in the general population.

Autistic traits remain moderately stable from childhood to adulthood (Taylor, Gillberg, Lichtenstein, & Lundström, 2017; Whitehouse, Hickey, & Ronald, 2011). Although autistic traits, like psychotic traits, are not viewed as problematic for psychological well-being in their milder forms (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), an increase in the number or severity of autistic traits is associated with increased anxiety (Kunihira, Senju, Dairoku, Wakabayashi, & Hasegawa, 2006), loneliness (Lamport & Zlomke, 2014), distress (Pisula, Kawa, Danielewicz, & Pisula, 2015) and shorter friendships (Jamil, Gragg, & DePape, 2017). In addition, Pisula, Danielewicz, Kawa, and Pisula (2015) found low to moderate negative correlations between autistic traits and aspects of quality of life.

As demonstrated above, there is ample evidence that the phenotypic traits related to ASD and psychosis found in the general population are attenuated versions of the same traits found in the clinical population. Taken together, this means both phenotypes are distributed along continua, ranging from clinical manifestations to phenotypic traits grading into normality.

**Similarities and Comorbidity.** In addition to the similar dimensional nature, SSD and ASD share much of the same characteristics. Both disorder spectrum have been linked to, for example, deficits in social cognition (Solomon et al., 2011), impaired recollection (Greenland-White, Ragland, Niendam, Ferrer, & Carter, 2017; Solomon, McCauley, Iosif, Carter, & Ragland, 2016), and language difficulties (Nunn & Peters, 2001; Solomon et al., 2011). The negative symptoms of SSD have also been compared to the impairments in social communication and social interaction in ASD (Hommer & Swedo, 2015). In addition, an unusual rate of comorbidity has recently been given more attention, which show that ASD and psychotic disorders seem to co-occur at higher rates than what would be expected in the general population.
Kincaid, Doris, Shannon, and Mulholland (2017) recently reviewed seven studies where participants diagnosed with psychotic disorders were screened for ASD or autistic traits. They found that the prevalence of ASD and autistic traits was higher in the clinical sample than one would expect in the general population. Solomon et al. (2011) reported that approximately 20% of patients with first episode psychosis (FEP) or clinical-high-risk for psychosis (CHR) met the criteria for Pervasive Developmental Disorder Not Otherwise Specified. When measuring for social responsiveness, 31% of the FEP group and 40% of the CHR group scored within the severe range, which indicates a strong likelihood of ASD.

A study by Barneveld et al. (2011) reported a higher rate of psychotic traits in ASD samples compared to community samples. Hallucinations and perceptual experiences have specifically been found to occur at higher rates in ASD samples compared to the general population (Milne, Dickinson, & Smith, 2017). Other studies have linked paranoid thoughts to ASD (Blackshaw, Kinderman, Hare, & Hatton, 2001; Jansch & Hare, 2014). Craig, Hatton, Craig, and Bentall (2004) found that people with Asperger’s syndrome reported higher rates of paranoia than typically developed adults, but less so than people with a delusional disorder. Interestingly, Unenge Hallerbäck, Lugnegård, and Gillberg (2012) examined the rates of ASD in multiple subgroups of SSD, and found that rates of ASD were considerably higher for individuals diagnosed with schizophrenia paranoid subtype. However, this link was not confirmed in the population-based study by Ford, Apputhurai, Meyer, and Crewther (2017).

There is also ample evidence that ASD and psychosis also overlap at trait level in the general population (Blain, Peterman, & Park, 2017; Gong, Wang, Lui, Cheung, & Chan, 2017; Sierro, Rossier, & Mohr, 2016). Ford et al. (2017) examined two separate scales, one measuring autistic traits and the other psychotic trait, and found several items from both scales which measured the same phenotypic traits in a non-clinical sample. These shared items were Social and Communication Discomfort, Odd Behaviour, Cue Interpretation and Relationship Disinterest, which the authors connect to the psychosocial symptoms that are the core features in both ASD and SSD.

**Theoretical Framework**

The unusually high comorbidity rate and the overlapping features warrant a theoretical model that can explain the shared characteristics in the autism and psychosis spectrums. One possibility is using a computational approach, which is an emerging approach within the field of psychiatry (Haker, Schneebeli, & Stephan, 2016; Huys, Maia, & Frank, 2016; Stephan & Mathys, 2014). The presumption is that brain is akin to a computing machine, where cognitions or thoughts can be compared to computations (Montague, Dolan, Friston, &
Dayan, 2012). Theory-driven approaches (mechanistic models with computational variables and relationships) or data-driven approaches (data-analysis using machine-learning techniques), or a combination of both, can be used to understand what phenomena underlies maladaptive cognition and clinical symptoms (Huys et al., 2016). One such theory-driven approach is the Bayesian approach (Huys et al., 2016).

The Bayesian approach builds on an earlier model of perception, which postulated that the brain is an inferences machine that updates sensory input in a generative model, creating a perceptual representations of the world which subsequently can be used to predict new sensory input (Dayan, Hinton, Neal, & Zemel, 1995). This model includes a generative model in the brain, which describes what input or data we could expect by generating synthetic data based on our previously learned sets of likelihoods. In other word, what you have previously learned about the world will tell you what you are likely to perceive next.

Generative models can also be used in a different context; that is, as models of the brain. These generative models are used to consider what data the brain would generate given all the possible ways a brain can function. One then looks at behaviour and infer which mental processes caused these. Generative models of the brain are especially important in computational psychiatry, as they can be used to investigate how maladaptive behaviour or symptoms are caused by latent, psychological processes (Stephan & Mathys, 2014).

In the generative model in the brain, we ask what we could expect to observe based on our knowledge of each state of the world (Haker et al., 2016). However, to ensure that our internal model is veridical, our beliefs must be examined in relation to observed data. In Bayesian inference, Bayes theorem is used to update the prior belief based on new data, creating a posterior belief. The posterior belief can subsequently be used as a new prior (belief) to predict new data, thus ensuring that a belief is continuously updated to reflect the real world.

An important feature of Bayesian inferences is the estimation of uncertainty. As our existing knowledge might be faulty or limited, it is necessary to include a subjective estimation of the how precise our beliefs are (Haker et al., 2016). Prior and posterior beliefs are represented as Gaussian probability distributions, where an increase in the width of the distribution (i.e. increased variance) reflects an increase in uncertainty. We perceive the belief to be more precise as the variance decreases. As the physical environment around us is constantly changing and our sensory organs are susceptible to noise, we also make an estimation of the precision of the data. These subjective precisions determine to what extent
the prior or the data is emphasised when the belief is updated. In graphical terms, the posterior probability distribution is weighted by the inverse variance of prior and data.

Importantly, these estimations of precision also affect the size (or precision) of the prediction error. A prediction error refers to the deviation between predicted data (based on prior belief) and the observed data. The size of the prediction error determines how informative a prediction error is, and whether it requires any attention. However, a prediction error might be unduly large due to either the data being perceived as more precise than it really is, or the prior being perceived as more precise than it really is. In Bayesian hierarchical models, prediction errors that cannot be reduced at lower levels becomes input for higher levels (Fletcher & Frith, 2009). The lower levels represent the cortical areas in the brain, and information becomes increasingly more abstracts as it ascends through the hierarchy (Haker et al., 2016). The higher levels represent the internal model of the world, which guides the information processing on lower levels. Thus, higher level beliefs influence sensory processing at lower levels, while lower level prediction errors can incite modifications of higher level beliefs.

C. Frith (2005) introduces the idea that delusions and hallucinations specific to schizophrenia are caused by false prediction errors in addition to an excessive tendency to attribute agency. False prediction errors refer to prediction errors that appear larger than they should be, thus demanding attention. The author bases his ideas on the notion that the sensory input caused by self-generated actions must be attenuated to make inferences about the world. This is because self-generated actions create sensory input in the same manner as input generated by external sources. For you to detect someone else’s breathing you must first discard the sound of your own breathing. As one can easily predict the outcome of self-generated actions, large prediction errors might indicate that the sensory input originated from an external source.

C. Frith (2005) argues that delusions of control can be explained by a greater tendency to attribute agency to unpredictable consequences. Fletcher and Frith (2009) extends this to all types of delusions, as all beliefs are changed when confronted with data that warrants an update. As a false prediction error is difficult to reduce, it is sent further and further up in the hierarchy, leading to a modification of the internal model at a higher and more abstract level. False prediction errors could therefore generate false beliefs with differing contents. Hallucinations can be explained by false prediction errors making (perhaps irrelevant) stimuli seem odd or surprising, accompanied by false beliefs from higher levels that guides sensory
experience. Thus, beliefs can influence the attention to stimuli, and how this stimuli is processed.

Similarly, Van de Cruys et al. (2014) argue that the symptoms of ASD can be explained by false prediction errors. Specifically, people with ASD give unduly attention to prediction error that in reality are so small that they should be ignored. It was previously mentioned that changes in the environment and noise from our sensory organs creates noisy data. The authors hypothesise that people with ASD fail to successfully distinguish between random variability and predictive cues, due to inflexible estimations of the size of the predictions error. Noise will make it difficult to match the new information to any higher-level representations, thus preventing new information to be generalised. Highly precise and rigid predictions could explain why the autism phenotype is associated with a preference for systems and patterns, as well as the reduced ability to generalise information (Van de Cruys et al., 2014).

Van de Cruys et al. (2014) explain how large and inflexible prediction errors can cause both social and non-social symptoms of ASD. Social interactions, which are complex and noisy, would require a lot of resources if all prediction errors are viewed as important. The abstract, higher-level representations would not be fully available to guide interpretations of facial expression, tone of voice, gestures, gaze and so on. This would explain why people with ASD find social interaction so difficult. Repetitive and stereotyped behaviour can be viewed as a way reduce prediction errors, by creating environments with highly predictive input. Fixating on highly restricted routines or arrays of interest can also be interpreted as a way of creating predictable environments, by focusing only on input that can be reliably predicted.

A Dimensional Model. According to the ideas of C. Frith (2005), Fletcher and Frith (2009) and Van de Cruys et al. (2014), false prediction errors cause the non-social and social symptoms in ASD as well as hallucinations and delusions in psychosis. How can the same underlying mechanism cause such different socio-cognitive symptoms? Recall that C. Frith (2005) attributed delusion of control to large prediction errors coupled with a tendency to attribute agency. Van de Cruys et al. (2014) do not discuss agency attribution in ASD, but the authors do refer to the mentalising problems in ASD as “mindblindness”, which refers to the reduced ability to represent mental states (U. Frith, 2001).

Agency attribution can be linked to mentalising, perhaps underpinning or acting as a precursor (C. D. Frith & Frith, 1999). In order to make judgments about someone’s mental state, one must first detect an animate agent with the capacity for mental states. Thus, if mindblindness in ASD reflects reduced agency attribution, the different expressions of false
prediction errors in psychosis and ASD can be explained by the phenotypes being positioned on opposite sides of a mentalising dimension. Abu-Akel et al. (2015) refer to this as the hypo- and hypermentalising dimension. Hypomentalising reflects a reduced tendency to attribute mental states, while hypermentalising reflect an excessive tendency to attribute mental states.

In the study by Abu-Akel et al. (2015), typically developed adults completed a digital perspective-taking task in which the participants followed a “director’s” instruction to move objects within a grid. To make the correct moves, the participants had to consider which object were visible to the director when he made the instructions. The authors found a relationship between high error rates and higher rates of autistic traits and psychotic traits. However, the combination of the two phenotypes showed a moderating effect. Specifically, the positive relationship between autistic traits and errors disappeared when these traits co-occurred with medium or high levels of psychosis traits. Similarly, the positive relationship between psychosis traits and errors disappeared when these traits co-occurred with high levels of autistic traits.

These results are similar to that of Shi et al. (2017). Their results showed that the interaction between autistic traits and psychosis proneness was negatively correlated with cognitive disability (inverse of executive functioning), while either autistic traits or psychosis proneness alone was positively correlated with cognitive disability. When considering total schizotypy scores, the interaction between autistic traits and schizotypy correlated positively with social functioning. This is in line with the ideas of Abu-Akel et al. (2015), who argues that if the phenotypes are positioned on opposite sides of a hypo- and hypermentalising dimension, the co-occurrence of traits from both phenotypes should buffer against the impairments associated with either one alone.

**Mentalising**

Several studies on social functioning and mentalising in ASD and psychosis indicate the existence of the hypo- and hypermentalising dimension proposed by Abu-Akel et al. (2015). Although there are studies that demonstrate the link between impaired mentalising and psychotic traits (e.g., Pickup, 2006) as well as autistic traits (e.g., Mintah & Parlow, 2018), only reports of directional errors (rather than errors in accuracy) can inform on the legitimacy of the hypo- and hypermentalising dimension. That is, one must look at research that measure the extent to which someone attributes mental states.

Non-verbal communication skills are essential for navigating the social world around us. We receive and convey information using movement, body language, gaze and facial expressions, and use this information to interact with other people. One can therefore look for
hypo- and hypermentalising behaviour in studies on, for example, body language and motion. A study conducted by van Boxtel, Peng, Su, and Lu (2017) showed a negative correlation between autistic traits and the participants’ ratings of the level of interactions of two stick figures. An increase of autistic traits was related to a greater tendency to label meaningful interaction as less interactive, as well as a reduced ability to distinguish between true social interactions and non-interactive motion. On the other hand, psychotic patients, specifically patients with delusions, demonstrates a tendency to interpret random or incidental movements as intentional gestures (Bucci, Startup, Wynn, Baker, & Lewin, 2008). This study also showed that patients with delusions tend to interpret gestures as insulting, despite the neutral content of the gestures.

The perception of facial emotions is another feature related to social perception, and therefore also to social functioning (Kaiser & Shiffrar, 2012). Perception of facial emotions involves the identification and interpretation of emotions in the facial expressions of others (Eack, Mazefsky, & Minshe, 2015). A noteworthy study by Pinkham, Brensinger, Kohler, Gur, and Gur (2011) showed that non-paranoid and paranoid patients with schizophrenia did not differ in the ability to accurately recognize facial expression, yet only paranoid patients demonstrated a bias. Specifically, patients with paranoid delusions tended to attribute anger to neutral faces.

Another study with patients diagnosed with first episode psychosis also showed a tendency to misidentify neutral faces (Catalan et al., 2016). Meta-analyses show a robust deficit in facial emotion recognition for people with schizophrenia, but varying in effect sizes and potential moderators (Chan, Li, Cheung, & Gong, 2010; Kohler, Walker, Martin, Healey, & Moberg, 2010). Catalan et al. (2016) suggested that atypical perception of facial emotions might be a trait of psychosis, which means that the same deficits would be visible at the sub-clinical end of the continuum.

This over-attribution of emotion to neutral faces have also been found in people with ASD (Eack et al., 2015). However, in this case, people with ASD showed the tendency to both over-attribute emotion to neutral faces and under-attribute emotional faces as neutral, suggesting a general confusion with neutral facial expressions (Eack et al., 2015).

In addition to the behavioural studies mentioned above, there is also evidence from the field of neuroscience. One example is Backasch et al. (2013), who investigated patterns in neural activation when presenting videos depicting cooperative (two people manipulating an object together) and non-cooperative (one person manipulating an object while the other watched) situations. When presented with non-cooperative videos, patients with delusions
showed more activation in brain networks associated with mentalising processing compared to the typically developed individuals. Importantly, both positive symptoms in the clinical group and the analogous cognitive-perceptual traits in the control group correlated positively with activation of the same neural networks in non-cooperative trials. This supports the dimensional aspect of hypermentalising behaviour.

Similar studies have shown that psychotic and autistic traits are linked to diametric patterns of activations in brain areas associated with mentalising, such as the right posterior superior temporal sulcus and the ventral medial prefrontal cortex (Ciaramidaro et al., 2015), and the posterior and anterior subdivisions of the ventral right temporoparietal junction (Abu-Akel, Apperly, Wood, & Hansen, 2017). Ciaramidaro et al. (2015) found that hypomentalising in autism was related to less activation for processing information concerning mental states, while hypermentalising in psychosis was related to increased activation for processing information concerning physical causality. The authors argued that this demonstrated that physical causality is proceed in a similar manner as mental states in psychosis, and that this might suggest that psychosis is related to a hyperactive detection of intentionality.

Agency Attribution

As we have seen, there is evidence that people on the psychosis spectrum and people on the autism spectrum show impaired mentalising abilities but reflecting opposite sides of the hypo- and hypermentalising dimension. Tasks measuring mentalising usually contain cues that indicate the presence of an agent (Blakemore, Sarfati, Bazin, & Decety, 2003). For example, the detection of biological motion might be a precursor to attributing emotion to motion (Nackaerts et al., 2012). Conversely, making inferences about mental states indicates the detection of an agent with the capacity for mental states. Thus, this hypo- and hypermentalising dimension should also be evident in research on agency attribution.

The investigation of the attributional styles associated with delusions and Asperger’s syndrome suggests that the phenotypes are distinct in how they attribute the cause of events (Craig et al., 2004). The delusional patients made more external-personal attributions directed at outside agents (e.g., “He did not call because he does not like me”), while patients with Asperger’s syndrome made more situational attribution (e.g., “He did not call me because his phone had no reception”). Delusional patients also tended to attribute the cause of negative events to external sources, while typically developed people attributed the cause of negative events more to themselves. This suggests the delusional patients have a greater tendency to attribute agency, and thereby responsibility, to outside agents. However, patients with
Asperger’s syndrome demonstrate the opposite pattern. One should note that these different attributional styles occurred despite patient groups getting similar scores on mentalising task, which supports the idea of an mentalising dimension.

The study conducted by Craig et al. (2004) might also explain the why people on the autism spectrum often report psychotic traits. As previously mentioned, the Asperger’s groups had a lower paranoia score than the delusional patients, but higher than the control group. The authors reasoned that the mechanisms contributing to paranoia must be different in patients with delusions and patients with Asperger’s syndrome, as the latter group did not show the attributional abnormalities that were found the delusional group. This argument fits nicely with the dimensional perspective, in which this mechanism is the tendency to attribute agency.

Gray, Jenkins, Heberlein, and Wegner (2011) examined the attribution of agency using a more nuanced conceptualization of agency, by distinguishing between agency (capacity for memory, self-control and acting morally) and experience (capacity for experiencing fear, pleasure and hunger). The results showed that people with more autistic traits tended to attribute less agency to adult humans, but the tendency to attribute experience did not differ. No relationship was found between autistic traits and agency or experience attribution to non-living humans, animals or objects. In contrast, there was no associated between attribution of agency to adult humans and scores reflecting cognitive-perceptual schizotypy. However, people with high schizotypy scores tended to attribute more agency and/or experience to non-living humans, animals and God. The results demonstrate that psychosis is linked to attributing mental capacities where there should be none, while autistic traits is related to attributing reduced mental capacity where there should be more.

A study on sense of agency also provides support for the hypothesised link between psychosis and attribution of agency (Maeda et al., 2012). Maeda et al. (2012) investigated perceptual causality in patients with schizophrenia paranoid type, by asking the patients to state whether they believed their self-generated action (pressing a button when hearing a sound cue) caused the outcome (a square to jumping on the screen). When the action preceded the outcome, positive symptoms correlated positively with the tendency to attribute the outcome to self-generated actions, while negative symptoms showed the inverse correlation. Remarkably, patients also demonstrated this positive bias when the outcome preceded the action. Patients with paranoid delusions were more likely than the control group to detect contingency between action and outcome, despite the square jumping before the button was pressed.
While the study above examines perceptual causality related to perceived agency of the self, animacy can be used to examine perceived agency of others. Both phenomena are irresistible in the sense that knowing that something is not animate or causal does not hinder the automatic perception of causality and animacy (Scholl & Tremoulet, 2000). Infants as young as 7-months show the ability to distinguish between animate and inanimate when presented with ambiguous shapes in motion (Träuble, Pauen, & Poulin-Dubois, 2014). Although one can argue that the link between sense of agency and agency attribution is not yet clear, the findings of Maeda et al. (2012) is advocated by similar results from an earlier study on animacy using geometric shapes (Blakemore et al., 2003).

Blakemore et al. (2003) investigated the perception of contingency using shapes with animate and inanimate movements. The authors applied a 2x2 design by making the two shapes to display both contingent relationships (the movements of first shape causing the movements of the second shape) and non-contingent relationships (the movements of the second shape was not related to the movements of the first shape). They then asked patients with and without persecutory delusions, as well as typically developed individuals, to rate the strength of the relationship between the shapes.

When presented with shapes with inanimate movements, all participants rated the shapes in the contingent condition as having stronger relationships than those in the non-contingent condition. When viewing animate shapes, non-delusional participants, both patients and controls, perceived a similarly strong difference between contingent and non-contingent movement. However, patients with delusions did not consider the relationship in non-contingent condition to be different from the relationship in the contingent condition. Some of the delusional patients argued that the second shape in the could hear or feel the first shape. This demonstrates the over-attribution of agency, in which a non-existing interaction between geometric shapes is anthropomorphised to the extent that one shape’s independent movements are perceived to be caused by an external agent.

Blakemore et al.’s (2003) findings corresponds to the results from a study using point-light stimuli, in which paranormal beliefs was measured in the general population (Elk, 2013). When the author presented the participants with stimuli containing low to intermediate ambiguity, increased paranormal beliefs was related to a higher tendency to attribute agency. The author attributed the non-significant relationship in highly ambiguous trials to the nature of the stimuli; excessive agency attribution only occurs when the stimuli can be perceived as meaningful. This study augments studies like those of Blakemore et al. (2003) and Maeda et
al. (2012), by demonstrating the psychotic traits in the general population also can be linked to excessive agency attribution.

Problems and Gaps

Artificial stimuli in motion, like the geometric shapes in the study by Blakemore et al. (2003), are well suited to study animacy. They are more easily manipulated and less ambiguous than complex, naturalistic stimuli, yet readily perceived as animate. Indeed, by manipulating factors such as speed and trajectory, a single white dot moving across a dark background appear animate to observers (Tremoulet & Feldman, 2000). Therefore, using animated stimuli to induce the perception of agency might be a more reliable way to measure agency attribution than using self-report (e.g., Gray et al., 2011).

Stimuli in motion also covey social information, like emotional states (Clarke, Bradshaw, Field, Hampson, & Rose, 2005) and intentional interactions (Barrett, Todd, Miller, & Blythe, 2005). Moreover, Miller and Saygin (2013) used point-light displays of walkers to demonstrate that individual differences in social cognition influence the perception of animacy. When observers interpreted biological motion using global information (i.e. focusing on all moving dots rather than individuals or groups of dots), better mentalising abilities was related to better task performance.

However, concerning the study of social perception and social cognition, one might argue that artificial stimuli might be too limited in terms of ecological validity. While there exists evidence that observers readily perceive agency in moving dots, it is not clear whether one could manipulate the same features (like speed, shape and trajectory) to elicit mental state attribution. Although the hypo- and hypermentalising dimension seems to be supported by some research on agency attribution, using animated shapes to measure attribution of higher order mental states is more controversial. This applies, for example, to the use of the animations of Heider and Simmel (1944), in which interacting triangles reportedly behave as if possessing representations of the mental states of others.

Stimuli material building on Heider and Simmel’s animations have been used in several research studies, replicating the phenomena in which the triangles are perceived as interacting agents (for review, see, Scholl & Tremoulet, 2000). Some of these studies appear highly relevant for the current topic, especially those studies which also include triangles with less “social” movements for comparison. One example is the study by Russell, Reynaud, Herba, Morris, and Corcoran (2006), where patients with paranoid delusion and patients with passivity experiences described randomly moving triangles as interacting or even referring to mental states, in line with hypermentalising proposition (Abu-Akel et al., 2015).
More important is the study by Castelli, Frith, Happe, and Frith (2002), who used the same animations to study mental state attribution in autism. The authors asked typically developed people and people with autistic disorder and Asperger’s syndrome to describe the movements of the triangles, and then rated those descriptions based on intentionality (agency) and appropriateness. The groups differed only when presented with triangles depicting theory of mind behaviour (i.e. persuading, mocking, bluffing and surprising), where the patient group used fewer mentalising descriptions. Importantly, they did not differ when presented with animations performing goal-directed actions (i.e. dancing, chasing, leading and imitating).

If under-attribution of agency was part of or underpinning hypomentalising behaviour in autism, then one might expect this to be reflected in lower agency ratings for goal-directed animations. If the goal-directed animations measured agency attribution, these results imply that hypomentalising associated with the autism phenotype does not extend to agency attribution. Findings like these should undoubtedly be part of the discussion, but they are not considered conclusive for several reasons. Most importantly, it is not clear that what Castelli et al. (2002) calls “goal-directed actions” is equivalent to identification of agency or intention, although the term “goal-directed behaviour” has previously been used to refer to the detection of intentionality (Tremoulet & Feldman, 2000).

The animations were assigned to the conditions (theory of mind, goal-directed, random) based on consensus in a pilot study of unknown size and composition (Abell, Happé, & Frith, 2000). The fact that animations were assigned to conditions without a scientific rationale challenges the validity of the results. It is not certain that what separates theory of mind animations from goal-directed animations is truly the elicitation of mentalising. Likewise, it is not evident that the difference between goal-directed and random movements is that only goal-directed movements elicit agency attribution or the perception of animacy. In fact, one of the goal-directed animations was pulled from the aforementioned study, as typically developed people tended to interpret the animation differently from the consensus reached in the original study (Russell et al., 2006).

It is also worth noting that this task require the participants to produce verbal descriptions. As previously mentioned, language impairments have been linked to both psychosis and ASD. Castelli et al. (2002) stated that the groups did not differ in verbal ability based on the results of the Ammons Quick test (Ammons & Ammons, 1962). However, this is a general intelligence measurement that does not include any speaking or writing. If the language skills in the patient group are only slightly below those in the control group, one could expect the group effect to only be visible in instances where the control group produce
high intentionality ratings. Taken together with the fact that the scoring procedures are based subjective interpretations, these results are judged too ambiguous to answer the research questions of the current study.

The use of clinical samples is a reoccurring feature of research on psychosis and ASD in relation to agency attribution and mentalising which should be addressed. Using clinical samples is problematic for two reasons: First, behavioural data might be unreliable due to factors such as general impairment and medication. Second, samples that may appear similar might have been drawn from different populations. As previously mentioned, there is evidence that the phenotypic traits related to ASD and psychosis in the general population are the same traits that are measured on the clinical end of the continuum. The threshold that categorically separates healthy from clinical might vary when comparing studies using clinical samples. For example, while some studies on ASD include individuals with Asperger’s syndrome (e.g., Pellicano, Rhodes, & Calder, 2013), others do not (e.g., Senju, Tojo, Yaguchi, & Hasegawa, 2005). Findings from clinical studies must therefore be interpreted with caution.

The present study

In sum, there are findings that suggests the hypo- and hypermentalising dimension proposed by Abu-Akel et al. (2015) extend to agency attribution. Yet, there is a lack of conclusive evidence that describes how agency attribution relates to the tendencies towards ASD and psychosis. The present study aims to investigate the hypo-and hypermentalising dimension by examining how the traits relate to agency attribution in a detection task designed to evoke the perception of animacy. The rationale behind this study is based on theory and literature that suggest mentalising extends to agency attribution, and that animated stimuli readily and reliably induce agency attribution. By examining detection, rather than ratings or descriptions, one can calculate whether participants are generally biased to see or not see agency, both when agency is present and absent.

No other study has previously examined both autistic and psychotic traits within the same agency detection paradigm. This is an essential step to take to infer whether Abu-Akel et al.’s (2015) dimensional model can be applied to the Bayesian theories of psychosis and ASD. In other words, this study might offer findings that can be used to explain the behavioural differences associated with psychosis and ASD.

The present study uses animations based on the wolf and sheep task of Gao, Newman, and Scholl (2009), which was designed to elicit the perception of chasing. The authors argued that the perception of chasing is a salient and robust form of animacy, perhaps carrying
evolutionary advantages (Gao & Scholl, 2011). The simplicity of the task makes the wolf and sheep task appropriate for the current sample. The task does not require the participants to give any verbal or written responses, nor does it include any subjective scoring formats. However, as the material is modified for the present study, the stimuli is considered novel. A measure of anthropomorphism will therefore be used for comparison.

Like animacy, anthropomorphism, the tendency to attribute human characteristics, goals and intentions to non-human entities, is also a process that happens automatically and readily (Epley, Waytz, & Cacioppo, 2007; Waytz, Cacioppo, & Epley, 2010). The individual differences in the tendency to anthropomorphise appear to remain reasonably stable of time, suggesting that the tendency reflects a behavioural trait (Waytz et al., 2010). Waytz et al. (2010) argues that anthropomorphising extends beyond agency attribution and animacy, as it also entails attribution of mental states to living entities (like animals). However, as anthropomorphising non-living entities entails perceiving these as agents capable of intentional actions, measures of anthropomorphising has previously been used as a proxy for measuring agency attribution (Imhoff & Bruder, 2014).

The first objective of this study is to investigate how bias to see agency relates to the tendencies towards autism and psychosis. The second objective is to investigate whether the individual’s tendency to anthropomorphize predict bias to see agency in the wolf and sheep task. The most crucial measure from the wolf and sheep task is therefore the individual’s tendency to detect chasing (as measured by response bias $c$). In addition, the ability to detect a signal (as measured by $d'$) informs us about the difficulty level of the wolf and sheep task. The $d'$ shall therefore be used to interpret the appropriateness of the stimuli.

As the phenotypic traits distributed in the general population reflects the same traits found in the clinical population, this study will measure traits in a general population sample. Using a non-clinical sample is also important to investigate the effects of traits, as general cognitive impairment and medication in clinical samples can obscure any effect found.

The present study was pre-registered on the Open Science Framework (https://osf.io/), which means that the hypotheses, sampling plan, design and analysis plan was determined before the data collection started. The preregistration was approved November 11th, 2017, and is available online, along with data and pilot data, at https://osf.io/n3rbp/. It is important to clarify that because these hypotheses were preregistered, no amendments have been made after the approval date when contradictory findings have emerged in the literature. This is, however, only relevant regarding hypothesis five. This hypothesis and relevant findings will
be further discussed in the discussion. The following hypotheses were preregistered and approved for this study:

Hypothesis one: Anthropomorphism (measured by the Individual Differences in Anthropomorphism Questionnaire [IDAQ]) and bias to see a wolf in the wolf and sheep task measure the same construct and accordingly correlate positively. In all following hypotheses, it is assumed that anthropomorphism and wolf and sheep bias have the same relationship to other variables. If that correlation is low, they must be treated as separate predictors in all the following hypotheses.

Hypothesis two: Anthropomorphism and bias to see a wolf correlate positively with the positive symptoms of psychosis, as measured by the positive subscale of the Community Assessment of Psychic Experiences (CAPE 42).

Hypothesis three: Existing research implies conflicting predictions regarding the relationship between, on the one hand, anthropomorphism and bias to see a wolf and, on the other hand, tendencies towards autism (as measured by Autism Spectrum Quotient [AQ]). A) If anthropomorphism and bias correspond to the hypo- versus hypermentalising dimension proposed by Abu-Akel et al. (2015), then anthropomorphism and bias to see wolf should correlate negatively with AQ. B) If instead the attribution of intention I hope to measure with anthropomorphism and bias corresponds to detection of goal-directed action, as opposed to identification of intention/agency, the results of Castelli et al. (2002) imply that anthropomorphism and bias should be uncorrelated with AQ. C) Several studies show elevated incidence of paranoid ideation associated with ASD (see above), which would suggest that anthropomorphism and bias to see a wolf correlate positively with AQ. I therefore have no directional prediction.

Hypothesis four: Anthropomorphism and bias to see a wolf do not correlate with the ability to distinguish trials with a wolf from trials without in the wolf and sheep task (i.e. the ability to detect signal, as measured by d’).

Hypothesis five: Tendencies towards autism does not correlate with tendencies towards psychosis (I want to remind the reader that I found papers indicating a positive correlation only after this hypothesis had been preregistered. This hypothesis is still included as this list of hypotheses is the one I committed to in the preregistration. See above).

Hypothesis six: Tendencies towards autism do not correlate with ability to distinguish trials with a wolf from trials without in the wolf and sheep task.

Hypothesis seven: When it comes to the correlation between tendencies towards psychosis and the ability to distinguish trials with a wolf from trials without in the wolf and
sheep task, existing research conflicts with theory. A) Theory predicts that positive symptoms should not affect ability, which would mean no correlation between ability and CAPE 42 Positive Subscale. B) Roux, Passerieux, and Ramus (2015) found a decreased ability in a clinical sample with schizophrenic patients, which suggests that ability would be negatively correlated with the CAPE positive score. However, this study did not control for positive and negative symptoms. I therefore expect either a zero or negative correlation.

Hypothesis eight: I expect that ability to distinguish trials with and without wolf, correlates positively with the proportion of correct identifications of the wolf, and that the two variables have the same relationships to other variables. I therefore expect to replicate the relationships in hypotheses four, six and seven if I substitute ability to distinguish trials with a wolf from trials without by proportion correct identifications of the wolf.
Method

Participants

The participants were recruited by direct recruitments, by the distribution of an URL link on social media and at the Aspie Quiz website (http://rdos.net/eng/). Direct recruitment happened primarily at the Norwegian University of Science and Technology campuses Dragvoll and Gløshaugen, as well as at University of Tromsø the Arctic University of Norway. Participation was voluntary, and no incentives were given. However, to ensure that most participants completed the study, they were told that their scores would be available to them upon completion. Participation was also anonymous, which means that no identifiable (e.g., names) nor personal information (e.g., IP addresses) was collected. As a result, the study was not subject to notification according to the procedures of Norwegian Centre for Research Data.

932 participants consented to participate.¹ Of these, the 553 (~59%) participants that failed to respond to 2/3 of each of the study’s indexes was excluded. Another 214 participants were excluded due to either being under the age of 18 years old, scoring an average of ≥ 2 on the CAPE 42 Positive Subscale control questions or indicating that they watched videos (see Material and Procedures) twice on more than 1/3 of the trials. Spending more than 12 seconds on the webpages that contain each video was interpreted as seeing the videos twice. Thus, the study’s sample included 165 adults (94 females, four participants identifying as “other”, and two participants did not disclose gender). The age of the participants ranged from 18 to 69, with a mean age of 31 years (SD = 10.8).

An additional exclusion criterion was part of our original preregistration, but it was later removed as its use was inappropriate for our data. The initial idea was to screen for participants who answered questionnaire items without reading the questions by excluding anyone who gave the same response to 1/3 of any questionnaire. An amendment to this criterion was registered prior to the data analysis to not include the CAPE 42 positive subscale, as results from previous studies indicated that the distribution would be skewed towards lower scores (Gillespie, Mitchell, & Abu-Akel, 2017; Wigman et al., 2017). The data screening revealed that scores from the Individual Differences in Anthropomorphising Questionnaire were also skewed toward the lower end of the scale, which, in the case of only the lowest response alternative, resulted in 141 of 165 participants being excluded. As the responses to the Autism Spectrum Quotient were scored on a 4-point Likert scale, a

¹ For more details on screening procedure, got to https://osf.io/n3rhp/.
comparison of responses across the four values revealed that 134 of 165 participants should be excluded. Thus, this exclusion criterion was abandoned.

**Measures and Procedure**

The participants completed an online survey composed of self-report measures for autistic traits, psychotic traits and anthropomorphism, as well as a visual search task assessing the perception of chasing. The survey was created using Qualtrics (https://www.qualtrics.com). The survey consisted of a total of 66 questionnaire items and 50 trials from the visual search task, which were divided into blocks with items and blocks with trials. To avoid fatigue, blocks with items and blocks with trials were presented in an alternating order. The items of each questionnaire were evenly divided into either two or three blocks, with a maximum number of 12 items per block. All questionnaire items were presented in their original order within each block. The trials were divided into blocks of seven, with one additional trial in the last block. The block order was fixed, but videos were randomized within each block. Both English and Norwegian versions of the survey was published online.

**Demographic Measures.** Age and gender were the only demographic variables collected, with the intention of preserving the participants’ anonymity (see above).

**The Individual Differences in Anthropomorphising Questionnaire.** The participants completed the 15-item version of the Individual Differences in Anthropomorphising Questionnaire (IDAQ; Waytz et al., 2010), assessing the participants’ tendencies to anthropomorphise non-human entities such as animals and technology. The participants used an 11-point scale (0 = not at all to 10 = very much) to rate the extent to which they believed the nonhuman entities possessed human characteristics, such as free will or emotions. The total scores range from a minimum of 0 to a maximum of 150, with higher number indicating greater tendency to anthropomorphise. The 15 non-anthropomorphic items from the original questionnaire were not included, as they do not measure anthropomorphism.

Support have been found for the questionnaire’s two-factor structure by using an independent sample (Cullen, Kanai, Bahrami, & Rees, 2014), and the IDAQ demonstrates good internal consistency (Letheren, Martin, & Jin, 2017) and moderately high reliability (Waytz et al., 2010). The Cronbach’s alpha of $\alpha = .824$ indicate a high level of internal consistency for the current sample.

**The Autism Spectrum Quotient – Short form.** The participants completed the abridged version of the Autism Spectrum Quotient (AQ; Hoekstra et al., 2011). This is a shorter version of the Autism Spectrum Quotient by Baron-Cohen et al. (2001). The short AQ
have shown to be highly correlated with the original 50-item questionnaire (Hoekstra et al., 2011). This questionnaire consists of 28 items assessing autistic traits along a continuum in the general population. The items were rated on a four-point scale (1 = definitely agree to 4 = definitely disagree). Traits are assessed along the five dimensions of social skills, routine, imagination, switching and numbers/patterns. The scoring procedures followed the four-response scoring format of Hoekstra et al. (2011), based on the recommendation by Kuenssberg, Murray, Booth, and McKenzie (2014). The total scores therefore range from 28 to 112, where the maximum score representing full endorsement of all autistic traits.

The internal consistency and reliability of the AQ-S has been validated using several independent samples (Hoekstra et al., 2011). In addition, a study by Kuenssberg et al. (2014) suggests that the proposed structure of the short AQ include structural and scale properties that can be extended to samples of individuals with the clinical diagnosis of Autism Spectrum Disorder. The Cronbach’s alpha of $\alpha = .926$ indicate a very high level of internal consistency for the current sample.

The Positive Subscale of the Community Assessment of Psychic Experiences. The participants completed the positive subscale of the Community Assessment of Psychic Experiences (CAPEpos; Stefanis et al., 2002). The self-report questionnaire consists of 20 items measuring the presences of positive symptoms, such as paranormal beliefs and bizarre perceptual experiences. The distress items were not included in the current study. However, three control questions reflecting common misconceptions about psychosis were added (Moritz et al., 2013). One the original control questions was not included due to its content being interpreted as not culturally relevant for the current sample. The items were rated on a four-point scale (0 = never to 3 = always) for frequency. Following the guidelines for scoring the Community Assessment of Psychic Experiences, the scores were recoded from 0-3 to 1-4.² The total scores range from a minimum of 20 to a maximum of 60, with the maximum score representing full endorsement of all symptoms of psychosis.

A study by Konings, Bak, Hanssen, van Os, and Krabbendam (2006) demonstrated good reliability and validity of the questionnaire, and the positive subscale of the CAPE 42 has previously been used to measure psychotic traits in general population samples (Booij, Snippe, Jeronimus, Wichers, & Wigman, 2018), clinical samples (Krkovic, Moritz, & Lincoln, 2017) and student samples (Capra, Kavanagh, Hides, & Scott, 2013). The

Cronbach’s alpha of $\alpha = .915$ indicate a very high level of internal consistency for the current sample.

**Visual Search Task.** The wolf and sheep task by Gao et al. (2009), which assesses perception of chasing, was modified for this study. New videos were created using custom software, which was written using the 16.0.0 version of LabVIEW (2016). These videos were dissimilar to the videos used in Gao et al. (2009) in several ways. This was the result of bringing overall performance to a level in which both ceiling effects and floor effects were avoided, thereby also ensuring that the participants would produce enough errors to calculate a meaningful bias $c$. This was achieved by modifying the parameters (video length, speed, number of distractors and chase/escape angle) during pilot testing to produce a $d'$ as close to 1.0 as possible. The parameters described in the following sections were thus chosen based on their contribution to the optimal difficulty level.

The participants were presented with a total of 50 videos. The videos contained 16 red distractor discs and one yellow disc moving in random directions, as if moving of their own accords. In half of the 50 videos, one of the 16 red discs would chase the yellow disc. The participants were presented with the cover story that this red disc was a wolf that was trying to catch the sheep (yellow disc). The yellow disc is hereafter referred to as the sheep, while the chasing disc is referred to as the wolf. The large number of distractors is a notable modification of the visual search task of Gao et al. (2009), who used an invisible sheep and a maximum of five distractors. In addition, the length of each video was reduced from 10 to approximately 5 seconds. All red discs were marked with a different letter, which were selected based on how easily one could discriminate between the letters. The participants were instructed to watch the videos once, and after each video they were asked, “was the yellow sheep being hunted?” In trials where the participants answered yes, they were asked to identify the wolf by clicking on a letter.

Before starting the trials, the participants were presented with two warm-up trials with decreased difficulty level, achieved by reducing the number of distractors. This method for decreasing the difficulty level was used to ensure that participants could clearly distinguish between warm-up trials and main trials, and adjust their strategies accordingly. Examples of videos, as well as data and pilot data, can be found online at https://osf.io/n3rbp/.

**Stimuli.** The discs, measuring at 34 pixels wide and 33 pixels long, moved across a black, 726x554 background. Discs were preferred over other shapes (e.g., arrowheads) as they do not reveal orientation, thus making it more difficult for the participants to detect the wolf (Gao & Scholl, 2011). The discs appeared in random location at the beginning of each video.
The discs moved in random directions, with each disc moving 12 steps of 5 pixels length in one direction before choosing a new direction randomly from \( \pm 90^\circ \) of the previous direction. This speed was constant for all discs. Each step was calculated every 20ms, so that a change in direction could be made if the discs encountered the edge of the frame before completing the 12 steps. The discs instantly chose a new direction when reaching the edge of the frame, by choosing a random direction from the \( 180^\circ \) range leading away from the edge.

In trials without chasing, the sheep moved the same way as other discs. In chasing-present trials, the sheep would choose a direction from the \( \pm 90^\circ \) range that was leading away from the wolf (fig. 1a). The wolf’s direction would be chosen from the \( \pm 90^\circ \) range leading towards the sheep (fig. 1b). Contrary to Gao et al. (2009), this chase angle was constant in all videos, as the results from pilot testing indicated that smaller angles created ceiling effects. Chase angles larger than \( \pm 90^\circ \) were discarded as they allowed for the wolf to choose a direction leading away from the sheep.

\[\text{Figure. 1.} \quad \text{An illustration of the chase angle and escape angle in chasing present trials. The grey areas show the angular zones from which new directions are chosen. The grey area, the lines and the arrows were not the part of the actual display. (a) The sheep (yellow disc) chooses a direction from the } \pm 90^\circ \text{ range (total grey area) leading away from the wolf (red disc). (b) The wolf chooses a direction from the } \pm 90^\circ \text{ range (total grey area) leading towards the sheep.}\]

\text{Statistical Analysis and Preliminary Analysis}

The d’ was within an acceptable range of our goal of d’=1.0 (see table 1), thereby ensuring that the wolf and sheep task yielded enough errors to calculate a meaningful bias c. The d’ reflects the performance level of the participants, where higher numbers represent better ability to detect when the wolf was present. Response bias c reflects the participants’
propensity to perceive the wolf as present or absent. A negative bias c score represents a bias to see wolf, while a positive score represents a bias to not see wolf.

As some participants did not complete the AQ, a proportional score was calculated to avoid biasing the AQ score of those with missing values. The range of possible values for the proportional AQ score range from a minimum of 1.0 to a maximum of 4.0, with an increase in value reflecting more autistic traits. Descriptive statistics, with an overview of mean scores of each index, are available in table 1.

Signal detection analysis was used to compute the relevant measures for the hypotheses. Hit rate and false alarm rate was transformed to z-scores to calculate response bias c and d’ from wolf and sheep task. Bias c and d’ was calculated according to Macmillan (1991) (see appendix 1). The proportion of trials in which the wolf was correctly identified (ID) was transformed using arcsine square root. This refers only to the proportion of trials in which the participants correctly reported seeing a wolf.

The hypotheses were examined by using Bayesian correlations. The rationale behind using Bayesian statistics rather than frequentists statistics is based on the fact that hypotheses four, five and six predict absence of relationships. As p-values only inform about the statistical chance of getting the results, they cannot be used to draw conclusion about whether there is evidence to support a null effect. In contrast, Bayes factor quantifies the support for a hypothesis by examining how the data can be explained by the null hypothesis compared the alternative hypothesis (or vice versa). Thus, as Bayes factor provides a basis to make conclusions about an absence of relationship. The Bayes factor was interpreted according to Wetzels and Wagenmakers (2012) descriptions. Thus, the threshold for evidence in favour of the alternative hypothesis is BF_{10} ≥ 3.0, while the evidence in favour of the null hypothesis (i.e. absence of relationship) is BF_{10} ≤ 1/3.

The normality test was calculated using IBM SPSS Statistics version 25, while correlations were performed using JAPS version 0.8.6 (JASP Team, 2018). Normality was tested using Shapiro-Wilk test, in which all variable violated the assumption of normality (p < .05), except the d’, p = .059. Bayesian Kendall’s tau-b was therefore chosen as the non-parametric correlation test. The correlations between CAPE 42 positive subscale and the IDAQ, CAPE 42 positive subscale and bias c, IDAQ and bias c, and d’ and ID was calculated using a one-tailed test, while the remaining correlations was calculated using a two-tailed test.
Results

Means and standard deviations are presented in Table 1, while all correlations are presented in Table 2. The first objective was to investigate whether response bias in wolf and sheep task predict tendency to anthropomorphise. Contrary to predictions, IDAQ was unrelated to bias c, \( r_{tb} = .040, \text{BF}_{10} = .061, \) one-tailed. This Bayes factor implies that the data is .061 times more likely to observed under the alternative hypothesis (i.e. a positive correlation) than the null hypothesis (i.e. an absence of correlation). This was interpreted as strong evidence for an absence of relationship between tendency to anthropomorphise and bias to see wolf. Thus, IDAQ and bias c were examined independently in relation to the remaining indices.

Table 1.

Descriptive statistics.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDAQ</td>
<td>33.2</td>
<td>18.8</td>
</tr>
<tr>
<td>CAPEpos</td>
<td>36.5</td>
<td>8.91</td>
</tr>
<tr>
<td>AQ</td>
<td>2.70</td>
<td>0.58</td>
</tr>
<tr>
<td>d'</td>
<td>1.45</td>
<td>0.64</td>
</tr>
<tr>
<td>Bias c</td>
<td>0.37</td>
<td>0.50</td>
</tr>
<tr>
<td>ID</td>
<td>1.32</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Note. IDAQ=Individual differences in anthropomorphism questionnaire, CAPEpos=CAPE 42 Positive subscale, AQ=Autism Quotient (Short form) proportional score, ID=Proportion of trials in which the wolf was correctly identified, transformed using arcsine square root. \( n = 165. \)

As predicted, the CAPE 42 positive subscale correlated negatively with bias c, \( r_{tb} = -.13, \text{BF}_{10} = 4.06, \) one-tailed, and positively with IDAQ, \( r_{tb} = .14, \text{BF}_{10} = 5.43, \) one-tailed. An increase in psychotic traits was related to an increased tendency to detect chasing in the wolf and sheep task, as well as an increased tendency to anthropomorphise. The Bayes factor indicated substantial evidence for both these relationships. Substantial evidence was also found for an absence of relationship between AQ and the IDAQ, \( r_{tb} = .063, \text{BF}_{10} = .21, \) two-tailed. On the other hand, the relationship between AQ and bias c was not meaningful, \( r_{tb} = -.089, \text{BF}_{10} = .42, \) two-tailed. Thus, only psychotic traits were related to
anthropomorphising and bias to see wolf, while the results imply that autistic traits were not related to the tendency to anthropomorphise.

The strongest relationship was, however, found when investigating the relationship between autistic and psychotic traits. The AQ was found to correlate positively with the CAPE 42 positive subscale, $r_{tb} = .46$, $BF_{10} > 100.0$, two-tailed. This Bayes factor represents decisive evidence for the positive relationship, in which an increase in autistic traits was related to an increase in psychotic traits.

The relationship between traits and $d'$ was examined to determine whether there was a relationship between traits and performance of the wolf and sheep task. Contrary to prediction, AQ was not found to be unrelated to $d'$, $r_{tb} = -.11$, $BF_{10} = .97$, two-tailed. The Bayes factor indicated no support for neither a relationship nor an absence of relationship between autistic traits and the ability to detect chasing. Similarly, the Bayes factor yielded no support for a relationship or absence of relationship between the CAPE 42 positive subscale and $d'$, $r_{tb} = -.099$, $BF_{10} = .59$, two-tailed. Thus, there were no meaningful relationships between the ability to detect chasing and the autistic and psychotic traits.

The relationship between the $d'$ and bias $c$, as well as $d'$ and IDAQ, was examined to investigate whether performance was related to bias to see wolf or tendency to anthropomorphise. Contrary to prediction, decisive evidence was found for a relationship between bias $c$ and $d'$, $r_{tb} = 0.24$, $BF_{10} > 100.0$, two-tailed. This positive correlation implies that an increase in ability to detect chasing was related to a stronger bias to not see the wolf. In contrast, there was substantial evidence for an absence of relationship between $d'$ and IDAQ, $r_{tb} = -.078$, $BF_{10} = .30$, two-tailed. As predicted, the ability to detect chasing was unrelated to the tendency to anthropomorphise.

Finally, the relationship between $d'$ and the proportion of correctly identified wolf (ID) was examined. As predicted, $d'$ correlated positively with ID, $r_{tb} = .36$, $BF_{10} > 100.0$, one-tailed, yielding decisive evidence for the hypothesis that an increase in ability to detect wolf was related to an increase in ability to identify the wolf. Accordingly, there was also decisive evidence for a positive correlation between bias $c$ and ID $r_{tb} = .21$, $BF_{10} > 100.0$, two-tailed, thereby replicating the relationships between bias $c$ and $d'$. As predicted, the relationships between $d'$ and traits were replicated in the relationships between ID and traits. ID was found to be unrelated to AQ, $r_{tb} = -.044$, $BF_{10} = .15$, two-tailed, while neither a relation nor absence of relationship was found between ID and CAPE 42 positive subscale, $r_{tb} = -.091$, $BF_{10} = .44$, two-tailed. However, the absence of relationship between $d'$ and IDAQ were not replicated by ID and IDAQ, $r_{tb} = -.093$, $BF_{10} = .47$, two-tailed. Thus, only the relationships between ability
to detect chasing and tendency to anthropomorphise was not replicated when substituting ability to detect chasing for proportion of correctly identified wolf.

Table 2.

*Bayesian Kendall’s Tau-b Correlation Matrix.*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. IDAQ</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. CAPEpos</td>
<td>.135</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BF$_{10} =$ 5.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(one-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. AQ</td>
<td>.063</td>
<td>.460</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BF$_{10} =$ .209</td>
<td>BF$_{10} &gt;$ 100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(two-tailed)</td>
<td>(two-tailed)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. d'</td>
<td>-.078</td>
<td>-.099</td>
<td>-.112</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BF$_{10} =$ .304</td>
<td>BF$_{10} =$ .593</td>
<td>BF$_{10} =$ .970</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(two-tailed)</td>
<td>(two-tailed)</td>
<td>(two-tailed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Bias c</td>
<td>.040</td>
<td>-.129</td>
<td>-.089</td>
<td>.241</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BF$_{10} =$ .061</td>
<td>BF$_{10} =$ 4.06</td>
<td>BF$_{10} =$ .421</td>
<td>BF$_{10} &gt;$ 100.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(one-tailed)</td>
<td>(one-tailed)</td>
<td>(two-tailed)</td>
<td>(two-tailed)</td>
<td></td>
</tr>
<tr>
<td>6. ID</td>
<td>-.093</td>
<td>-.091</td>
<td>-.044</td>
<td>.359</td>
<td>.210</td>
</tr>
<tr>
<td></td>
<td>BF$_{10} =$ .473</td>
<td>BF$_{10} =$ .444</td>
<td>BF$_{10} =$ .145</td>
<td>BF$_{10} &gt;$ 100.0</td>
<td>BF$_{10} &gt;$ 100.0</td>
</tr>
<tr>
<td></td>
<td>(two-tailed)</td>
<td>(two-tailed)</td>
<td>(two-tailed)</td>
<td>(one-tailed)</td>
<td>(two-tailed)</td>
</tr>
</tbody>
</table>

*Note.* IDAQ=Individual differences in anthropomorphism questionnaire, CAPEpos=CAPE 42 Positive subscale, AQ=Autism Quotient (Short form) proportional score, ID=Proportion of trials in which the wolf was correctly identified, transformed using arcsine square root.

$n = 165.$
Discussion

The aim of the present study was to investigate whether tendencies towards ASD and psychosis relates to bias to see agency in a chasing detection task. The results showed that psychotic traits were related to a bias to see agency, as indicated by the traits’ relationships to IDAQ and bias c. However, there was no clear association between autistic traits and bias to see agency, as indicated by the traits differing relationships to IDAQ and bias c. As the material was novel, a second aim of the study was to investigate whether the individuals’ tendency to anthropomorphize could predict bias to see agency in detection task. Contrary to prediction, the tendency to anthropomorphize did not predict bias to see agency.

The Relationship Between IDAQ And Bias c

This study demonstrated that bias to see agency in the wolf and sheep task was unrelated to the tendency to anthropomorphise. This suggests that the bias c and the IDAQ does not measure the same behaviour. As tendency to anthropomorphising previously has been used as a proxy for agency attribution, one might argue that this result implies that bias to see agency does not measure agency attribution. However, the perception of chasing has previously been found to reliably and readily induce the perception of animacy (Gao et al., 2009; Gao & Scholl, 2011). It might therefore be more reasonable to assume that the IDAQ does not measure agency attribution.

One possible interpretation is that the IDAQ taps into social judgement, while bias to see agency only involves social perception. Agency attribution and anthropomorphising seem similar in that they both happen automatically, and thus one might detect the capacity for mental states just as one detects animacy. However, the IDAQ required the participants to rate animals and objects on characteristics that are, perhaps, uniquely human. Waytz et al. (2010) argued that anthropomorphism extend beyond perceiving an agent as animate, because the agent is also attributed characteristics that are only associated with humanness. As we make top-down judgements or inferences about other humans’ mental states, it is plausible that attributing those same mental states to nonhumans in the absence of stimuli also involve top-down processes. The participants’ rating might therefore reflect top-down judgements that are not necessary for the detection of animacy cues. Comparing self-report measures to a behavioural task might also cause disparity. However, evidence of an absence of relationship might be better explained by anthropomorphism extending beyond the detection of agency.

It is important to note that although anthropomorphism might extend beyond agency attribution, it would not be appropriate to treat the IDAQ as proxy for mentalising. This refers
to both the ability to mentalise, as well as hypo- and hypermentalising tendencies. Anthropomorphising has previously been linked to brain areas associated with mentalising and action understanding (Cullen et al., 2014). However, the research on anthropomorphism has generally been focused on accuracy and functioning (Epley et al., 2007), rather than investigating the relationship between anthropomorphism and mentalising. It is therefore not clear how these two processes are related.

It should also be noted that there is ample evidence that suggests autistic traits is related to poorer mentalising ability and a hypomentalising tendency. If the IDAQ measured higher order mental state attribution, on level with commonly used mentalising tasks, then previous research implies that the AQ should correlate negatively with IDAQ. The fact that I found IDAQ to be unrelated to the AQ implies that IDAQ and mentalising tasks might measure different constructs.

**IDAQ, Bias c and Traits: Mental state attribution and Traits**

The absence of relationship between IDAQ and bias c implied that these indices would be examined separately with the remaining indices. Therefore, if a one-tailed test was used to investigate a relationship between bias c and an index, the same one-tailed test was used to investigate the same relationship between the same index and IDAQ. However, it should be specified that even though they were treated as separate predictors, a hypothesis is not accepted on the basis of only one of these. A predicted relationship between bias c and an index must be confirmed by an analogous relationship between the index and IDAQ to accept a hypothesis.

The rationale behind this is twofold. First, in the case of diametrically opposite findings, there is no clear principle to follow to judge whether a hypothesis should be accepted or rejected. For example, recall that autistic traits were predicted to either correlate negatively, positively or be unrelated to IDAQ and bias c. It would be unclear how one would evaluate this hypothesis if the one found evidence for a negative correlation with IDAQ, but no evidence for a correlation nor absence of correlation with bias c.

This brings us to the second reason; it is more meaningful to interpret the overall patterns when considering theory and the research objective. Recall that the first objective was to investigate how psychotic and autistic traits relates to bias to see agency. In the paragraph above, I argued that IDAQ might measure more than bias to see agency. However, this is speculation. As IDAQ and bias was unrelated, they may measure different forms of mental states attribution, but one cannot say for certain that bias to see agency is only measured by bias c. Thus, as it is possible to examine the first research objective by saying
that bias to see agency is indicated by either IDAQ or bias c, a stricter criterion is need. Only relationships that are confirmed by both IDAQ and bias c are therefore treated as conclusive with respect to the hypotheses. If one is interested in interpreting only one relationship as evidence for the hypotheses, one should consider correcting for multiple comparisons before drawing new conclusions.

**Psychotic traits.** Psychotic traits were found to be related to an increased tendency to see agency in the wolf and sheep task, as well as an increased tendency to anthropomorphise. These results correspond with C. Frith’s (2005) idea concerning excessive agency attribution in psychosis, as well as the previously mentioned studies linking psychosis to excessive agency- and mental state attribution. The fact that both relationships were predicted by an a priori hypothesis affords more legitimacy to the interpretation of these results as support for the hypothesised relationship between psychosis and agency attribution. Therefore, these results suggest that psychosis is indeed associated with hypermentalising tendencies, according to the hypo- and hypermentalising dimension of Abu-Akel et al. (2015).

The fact that both correlations displayed the hypothesised relationships is also a reason not to dismiss them as meaningless due to their sizes. These sizes can be characterised as small, yet the pattern is still meaningful and should therefore not be ignored. I would also hesitate to dismiss these sizes because of the nature of the subject. As several factors contributes to behaviour, one might argue that it would be naïve to expect strong correlations when doing behavioural research. These relationships, although small, are still meaningful in the sense that they could represent observable effects on behaviour.

An interesting inference can be made based on a larger pattern in the results. As previously discussed, the fact that autistic traits were unrelated to tendency to anthropomorphise suggests that IDAQ does not measure mentalising. At the same time, the fact that bias to see agency and IDAQ correlated positively with psychotic traits suggests that the IDAQ measures some form of mental states attribution. This might be different from agency attribution, as IDAQ and bias c were unrelated. If anthropomorphising is equivalent to neither mentalising nor agency attribution, this could imply that there are at least three dimensions to mental state attribution. If mentalising does not extend to agency attribution, then the analogous relationship between psychotic traits and bias c indicate that there might be two dimensions to agency attribution.

**Autistic traits.** Unlike psychotic traits, autistic traits were found to be unrelated to tendency to anthropomorphise. This absence of relationship was not replicated by autistic traits and bias to see wolf, nor was any other relationships. Recall that conflicting findings
indicated that autistic traits should either be a) negatively related, b) unrelated or c) positively related to anthropomorphism and bias to see wolf. As the IDAQ’s and bias’s relationships to autistic traits do not correspond, neither of these explanations are accepted.

Although these results cannot be used to confirm that autistic traits are unrelated to agency attribution, I argue that they can be used to further explain the findings of Castelli et al. (2002). Recall that people with ASD did not differ from typically developed people when attributing levels of intentionality to a group of animations that consisted of triangles that were dancing, leading, imitating and chasing. Let us assume that this absence of evidence reflects evidence of absence (although this must be confirmed using statistics like Bayesian). It would be unlikely that both IDAQ and bias c would correspond to the intentionality rating of these animations, as the present study found IDAQ and bias c to be unrelated. As I did not find evidence that conclusively showed an absence of relationship between AQ and bias c, it is plausible that the intentionality ratings correspond to the IDAQ, which was found unrelated to the AQ. This would then correspond to the absence of group effect.

There are additional reasons to think that the intentionality rating measured anthropomorphism, as opposed to just agency attribution. These animations depicted social interaction, and the ability to perform them might be perceived as a cue signalling the ability to infer about the other triangle’s state of mind. One could argue that attribution agency entails the attribution of a state of mind. In fact, these triangles showed chasing movements, which was thought to measure agency attribution in the present study. However, the chasing movements were categorised with other movements that might involve more than just agency attribution. If one triangle could dance with or lead the other triangle successfully, then those triangles must arguably be able to infer about the other triangles state of mind to coordinate their movements. Thus, the participants may have attributed mental state capabilities to the triangles that involved the ability to represent the other triangle’s state of mind, which extends beyond animism.

In addition, information about orientation and size might also have contributed to increased anthropomorphism of these triangles. Orientation can be perceived as gaze, and the large triangle might have been viewed as the parent of the small triangle. In contrast, the lack of these features in the wolf and sheep task could imply that the discs were not anthropomorphised to the same extent that the goal-directed triangles.

When placing the present findings in the context of the theoretical framework, one must ask why the results showed no support for the hypomentalising account. One reason might be due to sampling strategy, as the AQ was created, at least partially, by using a student sample.
(Baron-Cohen et al., 2001). The fact that the survey was available online resulted in a more diverse sample than expected, as indicated by age. One can therefore argue that the AQ might not be entirely appropriate for the current sample. The item “I would rather go to a library than to a party” is an example of an item where demographic variables such as age or employment could create interaction effects. This might also be a plausible explanation for why the relationships between autistic traits and bias c was meaningless, while the relationship between psychotic traits and bias c was not. In future research, one should consider including a control measure of the AQ, such as a measure of student status.

Finally, one should evaluate what the meaningless relationship between AQ and bias c implies. The benefit of using Bayes factor instead of p-values is exemplified by this finding; this Bayes factor correspond to a p-value larger than the standard .05 threshold, which might have resulted in the interpretation that autistic traits were unrelated to bias to see wolf. The fact that the Bayes factor met neither the 3.0 nor the 1/3 threshold implies that this conclusion would have been false. Instead, this shows that the current data is insensitive, and that more research is therefore needed. This might be related to the inappropriate use of the AQ, as the relationship between psychotic traits and bias c was meaningful. In addition, it might also be attributed to the test’s ability to detect true (or an absence of true) relationships, which is known as power. The results from AQ and bias c suggests that the wolf and sheep task was low in power.

The argument concerning low power is supported by the mean d’, which indicates that the overall performance level was high. Although the d’ was judged to be acceptable, it did not reach the optimal value of 1.0, which is thought to be a value in which both ceiling and floor effect are avoided. Thus, the wolf and sheep task was too easy. A lower d’ would also have created more extreme biases, perhaps increasing the sizes of several of the currently significant relationships. The rationale behind this will be discussed in the context of d’ and bias c.

An Alternative Explanation. When considering the power of tests, perhaps one should ask the question in a different way: Does the investigation of hypo- and hypermentalising tendencies call for tests with higher levels power? The results from the present study can be put in the context of similar findings from two previously mentioned studies, in which the overall pattern of results (including psychotic traits) were interpreted as support for a hypo- and hypermentalising dimension. The fact that these patterns of findings corresponds might indicate that hypo- and hypermentalising tendencies are difficult to detect.
When Ciaramidaro et al. (2015) found results similar to those in the present study, the authors provided an alternative explanation that incorporated both the positive relationship between psychotic traits and bias to see agency/anthropomorphism, as well as a lack of association between the latter and autistic traits. The authors argued that they found support for hypomentalising in ASD and hypermentalising in psychosis, but the tendency depended on the context: hypomentalising occurred when the capacity for mental states was present, while hypermentalising occurred when there was a lack of mental state capacities. These results also correspond with the findings of Gray et al. (2011).

This interpretation predicts several of the relationship from the present study. They also predict that if a measure of mental states attribution was included in the present study, increased autistic traits would be related to a reduced tendency to attribute mental states. Psychotic traits might not be related to mentalising, as the hypermentalising tendency would not produce false alarms when attributing mental states to humans. Future research on this topic should therefore investigate how autistic and psychotic traits relate to mental state and agency attribution when the targets of those attributions vary in mental state capacity. For example, one could extend the research of Gray et al. (2011), by examining if similar results can be produced without the use of self-report measures.

The subject of power of tests is one that naturally follows this explanation. The results from the Ciaramidaro et al. (2015), Gray et al. (2011) and the present study can be attributed to measurements with too little power to measure hypo- and hypermentalising tendencies. These tests must be able to detect hypermentalising when mental capacities are expected, and hypomentalising when mental capacities are not expected. Another way of looking at it is like this: It might be easier to detect hypermentalising when one should not mentalise, and it might be easier to detect hypomentalising when one should mentalise. To investigate whether low power is an issue, one should look for ceiling and floor effects. Unfortunately, the distribution of scores from Ciaramidaro et al. (2015) and Gray et al. (2011) are not available.

Finally, one must also acknowledge that behavioural tasks might generally be reduced in their power to detect any effects, as demonstrated by an fMRI study by Kuzmanovic et al. (2014). Kuzmanovic et al. (2014) asked typically developed people and patients diagnosed with high functioning autism to rate the extent to which they perceived the motion of two dots to be “person-like”. While the scores did not differ significantly between groups, higher animacy ratings in the control group resulted in more neural activation in networks associated with social processing compared to the patient group. The authors reasoned that the dissociation between task performance and neural activation was caused by patients relying
on physical properties to rate the stimuli, whilst failing to spontaneously interpret the dots’ movements as social cues.

**The Relationship Between Traits**

The results showed that autistic and psychotic traits were positively associated, which means that those with more autistic traits also tend to have more psychotic traits. The relationship was not hypothesised (see hypothesis five), but nevertheless, not unexpected. As the hypotheses were preregistered, this hypothesis was not amended when literature emerged suggesting autistic traits are positively associated with psychotic traits. In the general population, correlations between autistic and psychotic traits range from 0.40 (Blain et al., 2017) to 0.48 (Wakabayashi, Baron-Cohen, & Ashwin, 2012). Thus, this finding corresponds to previous studies that measured the relationship between autistic and psychotic traits.

**The Relationship between D’ and Proportion of Correct Identification**

Both d’ and the proportion of trials in which the wolf was correctly identified was hypothesised to measure the ability to detect wolf. This was confirmed by a positive relationship between d’ and ID, which means that those with a higher ability to detect the wolf were also better at identifying the wolf. As in the case of bias c and IDAQ, this requires a stricter criterion when evaluating how other indices relates to ability. Only relationships confirmed by both ID and bias c are therefore treated as conclusive.

It should be mentioned that misidentifications were not treated as false alarms, because misidentifications does not necessarily imply that no signals were detected. Recall that the perception of animacy is an automatic process. This means that that the speed of which we detect cues of animacy is likely to be quicker than the process of becoming consciously aware of these cues. However, those with better ability to detect the signal are more able to successfully discriminate between noise and signal, thereby increasing their likelihood of identifying the wolf.

**The Relationship between Ability and Bias c, and IDAQ**

Contrary to prediction, a positive relationship was found between d’ and bias c, which was confirmed by the analogues relationship between ID and bias c. This relationship indicates that better ability was related to a greater tendency to not see wolf. There was also substantial evidence for an absence of relationship between d’ and IDAQ. However, no evidence was found for either e relationships or absence of relationship between ID and IDAQ. Thus, it is not evident how the ability to detect wolf is related to the tendency to anthropomorphise.
Lynn and Barrett (2014) have offered a utility-based explanation for the unexpected relationship between ability to detect wolf and bias. However, the explanation of this relationship is challenging to understand by only relying on a descriptive account. I have therefore attached a figure from Linn and Barrett (2014, p. 17) in appendix B (figure 2) that should be used as a supplement to this description.

Lynn and Barrett (2014) demonstrated that an inverse relationship between d’ and bias c reflects an optimisation of decision making. The authors argued that people modulate their bias according to their ability, with the goal of optimising decision making and maximising their utility. They demonstrated that an extreme bias compensates for a low ability to detect the signal. Thus, those with extreme biases do not show impairment in both ability and bias, but rather a single impairment (low ability) that causes extreme biases. This also implies that those with a low d’ and lack of bias is impaired both in their ability to detect the signal and in their ability to adjust their bias to compensate for the low d’.

Imagine that either unequal base rates or payoffs create a decision environment that induces a bias to see wolf. This could be achieved by, for example, informing the participants that there are twice as many trials with wolf than without, or that they will only receive points for each time they detect when the wolf was present. In this environment, participants view misses as costlier than false alarms. It is therefore more beneficial for the participants to generally have a bias to see wolf. However, those participants with low ability to detect the wolf will make more mistakes than those with high ability. As one type of mistake is costlier than the other, an extreme bias ensures that those with low ability misses the wolf less by prioritising false alarms over misses. Thus, the participants with low ability and extreme biases will receive a higher reward over a series of trials than those with low ability and less extreme biases.

The relationship between bias and ability is illustrated in Lynn and Barrett’s figure (2014, p. 17). As bias is modulated based on ability, one can draw a line that represent the optimal bias with regard to each ability level. This line is called the Line of Optimal Response, which is the exponential line in the figure. This line also mirrors the correlation between bias c and d’ in their respective decision environments. Bias c is positively related to d’ when the environment that induce a bias to see agency (i.e. most participants have a negative bias c value) and bias c is negatively related to d’ when the environment induce a bias to not see agency (i.e. most participants have a positive bias c value).
Although Linn and Barrett (2014) demonstrate why we should expect bias $c$ and $d'$ to be related, their utility-based account conflict with the direction of the relationship between bias $c$ and $d'$ in the present study. Figure 3 shows that the data from the present study depicts a cone shape, where extreme biases are related to lower $d'$. When scores extend over the horizontal line (i.e. bias scores range from a negative value to a positive value), types of decision environment are determined by on what “side” most of the scores are. However, while most participants have positive bias $c$ scores, the relationship between $d'$ and bias $c'$ is not negative. Bias $c$ and $d'$ demonstrates a relationship that characterises an environment that should induce a bias to see wolf, yet the participants show a bias not to see wolf.

*Figure 3.* The relationship between bias $c$ and $d'$ scores from the present study. The dashed line marks the separation between a decision environments that induces a bias to see wolf (underneath the dashed line) and a decision environments that induces a bias to not see wolf (above the dashed line).

Regrettably, there is no literature nor theories that can explain these data. In addition, there are no obvious measurement errors related to the wolf and sheep task that could explain these findings. These results represent an oddity, and I therefore have no credible interpretation to offer. However, I recommend that one should investigate whether these
findings can be replicated. The results from replication studies will help determine whether this atypical relationship between d’ and bias c reflects something meaningful.

**ID, d’ and Traits: Ability and Traits**

**Psychotic traits.** Psychotic traits were hypothesised to be either unrelated or negatively related to ability to detect wolf. However, there was no evidence to support either a relationship or an absence of relationship between psychotic traits and d’. The investigation of ID and psychotic traits yielded the same results. Thus, these findings cannot be used to draw conclusions regarding the relationships between psychotic traits and ability to detect signal in wolf and sheep task. Instead, this lack of results might be another indication of low power of the detection task. If the high mean d’ yielded ceiling effects, then this result might be the product of low variation in high d’ scores.

One can speculate whether the hypothesised negative relationship is justified when comparing these results to the findings by Roux et al. (2015) in the context of Lynn and Barrett’s (2014) utility-based explanation. Roux et al. (2015) showed that psychotic patients had a lower d’ than the typically developed controls, but bias c did not differ between groups. When considered the literature demonstrating that psychotic traits extend from subclinical to clinical, the findings of the present study imply that the patients in should have shown a negative bias. If this is true, then this could indicate that the patients did not demonstrate more extreme biases because they failed to calibrate their bias to their d’.

This leaves the question of why the present study could not replicate the highly significant association between psychosis and d’ from Roux et al. (2015). Based on Lynn and Barrett’s (2014) explanation of how lower d’s creates more extreme biases, one could have expected that psychotic traits to be negatively related to d’. Thus, as a negative bias increases with psychotic traits, d’ decreased with psychotic traits. Aside from low power, another possibility is that lower d’ and failure to calibrate bias c are impairments that only becomes evident at clinical end of the spectrum. This is another possibility that should be investigated further.

**Autistic traits.** While ID was unrelated to autistic traits, there was no evidence that supported either a relationship or absence of relationship between autistic traits and d’. Thus, the hypothesised absence of relationship between autistic traits and ability to detect signal in the wolf and sheep task cannot be rejected or accepted. As with psychotic traits, this might be another indication of low test power and ceiling effects.
Limitations and Strengths

Limitations. Perhaps the most important limitation of this study is the high $d'$, which has previously been discussed in the context of several unexpected findings. Although pilot testing was done to determine the parameters of the stimuli, constraints on time and resources limited the number of pilot studies that were performed. As the high $d'$ might have created ceiling effects, the lack of variation in the scores may have resulted in a detection task with lower power. As indicated by Lynn and Barrett’s (2014) study, the lack of low $d'$ means that few participants would have to compensate for their low ability by more extreme scores. It should also be mentioned that a larger sample would also have acted as countermeasure to the low power of the wolf and sheep task. However, seeing that the mean $d'$ was high, a larger sample size would not most likely not have removed the problem.

Another feature that reduces the research quality is the fact that the participants had the ability to play the videos more than once. Using a computer program would have constrained the number of times the participants watched the videos, but this idea was abandoned as it would make recruitment more challenging. A screening criterion was introduced with the intention of excluding those who may have watched the videos twice by controlling for the time spent on the webpages. Unexpectedly, this resulted in many participants being excluded. Therefore, one can safely assume that the samples size would have been larger if the participants were only able to watch the videos once. Reducing the sample size might also resulted in certain characterises being over- or underrepresented in the sample.

A measurement of mentalising was not included as adding more items the survey would likely increase the dropout rate considerably. At that point in time, there was no quantitative measurement of mentalising tendencies available that had the appropriate length. By not including a measure of mentalising the study become limited in its ability to provide conclusive evidence that validates the relationship between hypo- and hypermentalising and agency attribution.

While several meaningful patterns have been found, the fact that no evidence was found for either a relationship or an absence of relationship in several instances weakens one’s ability to make statements about the implication of these relationships. This also applies to those relationships that were not confirmed by the analogous relationship.

Strengths. Despite these limitations, this study also has several strengths. Examining relationships rather than group differences corresponds to the fact that autistic and psychotic traits seems to be distributed continuously in the population. Investigating both autistic and psychotic traits within the same detection paradigm offers more validity to any speculations
regarding the hypo- and hypermentalising dimension. Sampling from a general population rather than a clinical one also means that general impairment and medication would not bias any results.

By preregistrating the study, the reader can be confident that the procedure and analysis have been performed in an unbiased fashion. This entails that a detailed description of hypotheses, data collection plan and analysis plan has been published online, thus ensuring complete transparency. This also increases the reproducibility of the results. The reader is therefore ensured that the results do not include any selective reporting nor that any analytic choice (e.g., removing outliers or transforming variables) have been made after the data was available to the author. The reader is offered evidence that no hypotheses have been changed after data collection. Publishing results and data online also contributes to greater causes in science. Open data allows more for research, as well as acting as countermeasure to the publishing bias in psychological journals.

This is a novel study as this stimuli material has not previously been used to examine the relationship between agency attribution and autistic and psychotic traits. The notion that that hypo- and hypermentalising depend on context has been further expanded on, by comparing results from three different studies (the present study included). Finally, this is the first study, to my knowledge, that compares individual differences in anthropomorphism to agency detection using an animacy task.

Implications and Future Research

Implications. The findings from the present study have, first and foremost, theoretical implications for the dimensional model. The results show that autistic and psychotic traits, as measured here, do not demonstrate opposite patterns in attribution of mental states, as measured by the wolf and sheep task and the IDAQ. As there are no opposite patterns, there is also no evidence that the cooccurrence of traits would results in a buffer effect rather than an additive effect on behaviour. If the hypo- and hypermentalising dimension truly extends to agency attribution, as suggested by several lines of research, then the present study does not find support for claim that the autism and psychosis phenotypes can be placed on opposite sides of a dimension.

This study does, however, provide evidence that support the hypermentalising account of psychosis. These findings confirm the notions of C. Frith’s (2005) regarding excessive agency attribution in psychosis. Concerning autistic traits, evidence was only found for an absence of relationship between autistic traits and the tendency to anthropomorphised. If these results represent a true absence of relationship between autistic traits and hypomentalising,
then these findings indicate the need for a re-evaluation of the hypo- and hypermentalising model to only include the attribution of higher order mental states.

The significance of this reframing becomes more obvious when viewing it the light of the Bayesian framework in which the dimensional model was inserted. Although the phenotypes might both stem from false prediction errors, the differences in behavioural outcome might not be caused be opposite patterns in agency attribution. It should be specified that these findings do not disprove opposite patterns in mental state attribution.

An absence of relationship was found that has implication for future studies on agency attribution. The absence of relationship between bias c and the IDAQ can be used to argue that tendency to anthropomorphise should not be used as a proxy for agency attribution. The present study used an adapted chasing detection task that has previously been found to induce agency attribution (Gao et al., 2009). If one wishes to create new material to measure agency attribution, this finding indicate that a measure of anthropomorphism cannot reliably be used to control for the perception of agency. However, this absence of relationship might also indicate that mental state attribution entails a third dimension in addition mentalising and agency attribution. This is indicated by the fact that psychotic traits showed the predicted relationships with IDAQ and bias c. Alternatively, if mentalising does not extend to agency attribution, there might be two dimensions to agency attribution.

Finally, the significance of this study is also based on the several research areas that have been identified, in which new research is needed. In the following section, I have chosen to present addition suggestions for future directions which arguably would offer valuable knowledge this field of research.

**Recommendations for Future Research.** Several limitations concerning research quality or research strategy have presented. However, the most crucial flaw is the low power indicated by the high d’. Replication studies with larger sample sizes and a more sensitive wolf and sheep task is necessary to confirm the findings of the present study. Any replication should therefore take measures to create a more demanding wolf and sheep task, in addition to using a sample size of appropriate size, as this study indicates that effects might be small. A lower d’ would create more errors, which creates more room for bias and inflates bias (as indicated by Lynn and Barrett, 2014). Thus, a more demanding task would be more sensitive to bias. A replication study should therefore allocate the appropriate amount of time and resources on pilot testing the stimuli. By creating a more sensitive measure, one could also explore the alternative interpretation discussed above. This would verify whether re-evaluating the hypo- and hypermentalising tendencies as conditioned on context is warranted.
The research objectives of the present study are based on the rationale that mentalising extends to agency attribution. One could argue that because autistic traits have been repeatedly associated with reduced mental state attribution, yet autistic traits did not show the same pattern with agency attribution, agency attribution and mental state attribution must be dissociated processes. However, the sampling strategy might have rendered the AQ inappropriate for the sample. For future research directions, I therefore recommend including a measure of mentalising. If the AQ is also included, the mentalising task would act as control measure for the AQ, as indicated by previous studies on autistic traits an mentalising.

It might be necessary to specify that measuring the ability to mentalise is not alone sufficient, as one also need to measure individual tendencies in mental state attribution. That is, instead of investigating how accurately you attribute mental states to other people, one investigates to what degree you attribute mental states. A simple of example of this distinction is the difference between “Is Sally thinking about Tommy?” (are there mental states?) and “Was Sally’s smile intentional?” (which mental states are there?). The most important challenge with the latter is that ceiling effects can easily arise. Thus, one must be able to detect variations among those who readily detects intention.

If one decides the extend the survey, one should consider removing the IDAQ to avoid fatigue. The present study showed that it is difficult to make meaningful comparisons between the self-report measure and the behavioural task. The relationship between anthropomorphism and agency attribution should be explored in a separate study. I have presented the idea that the IDAQ measures social judgements, while the detection of agency concerns social perception. One way to explore the relationship between these two is to investigate whether individual differences in perception corresponds to individual differences in judgement.

My recommendation would be to extend on the study by McCabe, Houser, Ryan, Smith, and Trouard (2001). These authors used a trust game where the participants were paired with another participant or a computer as decision makers. The first decision maker has two choices; either 1) choose the safe option where both decision makers get the same, low reward, or 2) give the second decision maker a high reward. In the latter, the second decision maker must choose between keeping the high reward to himself or giving ~45% to the first decisions maker. The authors found that the prefrontal cortex, which was active when the participants cooperated with other participants, was less active when the participants played against a computer. It would be interesting to investigate how these findings relates to tendencies towards autism and psychosis, as well as tendency to anthropomorphise. One
could ask questions like “do participants with reduced perception of cooperative decisions also show reduced tendency to anthropomorphise?”.
Conclusion

The autism and psychosis phenotypes have both been hypothesised to stem from false prediction errors. Different behavioural outcomes have recently been attributed to opposite tendencies in mental states attribution. Literature suggest that a hypo- and hypermentalising dimension extend to agency attribution, but there is a gap in the literature. While agency attribution has previously been investigated in the context of psychosis and ASD, this is the first study to my knowledge that investigates whether autistic and psychotic traits relates to distinct biases to see agency using a detection task. This study could not provide support for opposite tendencies in agency attribution associated with autistic and psychotic traits. However, a meaningful pattern of results appeared suggesting psychosis can be linked to hypermentalising tendencies. This warrants further research on the hypo- and hypermentalising dimension, using measures more sensitive to bias to investigate whether the absence of findings can be attribute to floor and ceiling effects.
References


First Episode Psychosis, Borderline Personality Disorder and Healthy Controls. *PLoS One, 11*(7), e0160056. doi:10.1371/journal.pone.0160056


Appendices
Appendix A

Calculation of Bias c and d’

Bias c and d’ was calculated according to Macmillan (1991):

\[ \text{bias c} = -0.5(z(\text{hit rate}) + z(\text{false alarm rate})) \]

\[ d' = z(\text{hit rate}) - z(\text{false alarm rate}) \]

Following the recommendations of Macmillan (1991), extremes scores were corrected by converting 0 to 1/(2N), and 1 to 1-1/(2N), where N is number of trials with wolf.
Appendix B
The Relationship Between Bias and Ability.

Figure 2. An illustration of how bias c varies as a function of d’. Dashed lines are Line of Optimal Response (LOR), which demonstrates the optimal bias according to level of d’ and decision environment. The red squares below the horizontal line reflects data in which the environments induced a bias to see the signal, while the grey dots above the horizontal line reflects data in which the environments induced a bias to not see the signal. Reprinted from “Utilizing” Signal Detection Theory”, by Lynn, S. K. & Barrett, L. F., 2014, Psychol Sci, 25(9), p. 17.
**Appendix C**
The Study's Info Text in Norwegian

**Å se intensjon**

**Bakgrunn og formål**
Vår forståelse av den sosiale verden er knyttet til vår oppfatning av intensjonell atferd. Ulikheter mellom individer kan komme av ulike grader av tilbøyelighet for å oppfatte en atferd som intensjonell. Vi ønsker å undersøke om slike ulikheter er knyttet til autistiske og psykotiske trekk. Vi baserer oss på teorien om at disse trekkene er nedarvet i den generelle populasjonen, og at det derfor er vanlig at folk flest vil ha noen av disse trekkene. Alle kan derfor delta på denne studien.

Dette er et samarbeidsprosjekt mellom Universitet i Tromsø - Norges Arktiske universitet og Norges teknisk-naturvitenskapelige universitet, og omfatter et forskningsprosjekt ved Institutt for Psykologi (UiT) og en masterstudie ved Instituttet for Psykologi (NTNU).

**Oppgave: Finn ulven**
I denne undersøkelsen vil du se 50 videoer som varer i 5 sekunder hver. Alle videoene inneholder 16 røde prikker og én gul prikk. Den gule prikken er en sau, og noen ganger er én av røde prikkene en ulv som jakter på sauen. Denne ulven jakter på sauen i ca. halvparten av videoene. Videoene kommer i tilfeldig rekkefølge, og det er din oppgave å undersøke i hvilke videoer sauen blir jaktet på.


**Frivillig og anonym deltakelse**

Når du trykker deg videre (>>/→) samtykker du i å delta i prosjektet, samt at du minst 18 år.

Dersom du har spørsmål til studien, ta kontakt med Robert Biegler (robert.biegler@ntnu.no), Gerit Pfuhl (gerit.pfuhl@uit.no) eller Rebekka Solvik Lisøy (rebekkal@stud.ntnu.no)
Appendix D
The Study’s Info Text in English

Seeing intention
Background and purpose
Our understanding of the social world is tied to our perception of intentional behaviour. Differences between individuals may occur due to differences in propensity to perceive behaviour as intentional. We want to investigate whether such differences are associated with autistic and psychotic traits. We base our research on the theory that these traits are inherited in the general population, and most people will therefore exhibit some of these traits. Thus, anyone can participate in this study.

This is a collaborative project between the University of Tromsø - Norway’s Arctic University and the Norwegian University of Science and Technology, and encompasses a research project at the Department of Psychology (UiT) and a master’s thesis at the Department of Psychology (NTNU).

Task: Spot the wolf
In this survey, you will be presented with 50 videos lasting 5 seconds each. All videos contain 16 red discs and one yellow disc. The yellow disc is a sheep, and sometimes one of the red discs is a wolf that is hunting the sheep. The wolf hunts the sheep in approximately half of the videos. The videos appear in a random order, and it is your task to find out in which videos the sheep is being hunted.

You are going to watch each video once, and after each video you are asked if the sheep was being hunted. You answer either “yes” or “no”. If you answer “yes”, you get an additional question about the identity of the wolf. All the red discs have a different letter on them, and you choose the letter you believe the wolf had. When you are done, you continue to the next video.

In addition, you will continuously get questions on how you perceive the world around you, with topics such as imagination, communication and free will. You will also receive two questions on gaming habits. The survey takes between 20 and 30 minutes to complete, and you will get your result at the end of the survey. For optimal picture, it is important that you use a stable internet connection, as well as a computer screen rather than a smartphone or a tablet.

Voluntary and anonymous participation
This survey is anonymous and all information will be treated confidentially. The result will be made public online when the project ends in May 2018, but all information is made anonymous and your data can not be linked to you. Participation is voluntary, and you can withdraw from the study at any time by exiting the website. The only prerequisite for participating is that you are at least 18 years old.

By pressing (>> / →) you agree to participate in the project, as well as confirming that you are at least 18 years old.

If you have any questions, contact Robert Biegler (robert.biegler@ntnu.no), Gerit Pfuhl (gerit.pfuhl@uit.no) or Rebekka Solvik Lisøy (rebekkal@stud.ntnu.no).
Appendix E
Instructions in the survey

Instructions presented before warm-up videos:
Norwegian survey: «Du skal først få to oppvarmingvideoer, slik at du får sett hvordan undersøkelsen fungerer. Disse vil ikke telle med på resultatet ditt. Du vil få tilbakemelding på oppvarmingen, men ikke på resten av videoene i undersøkelsen.»

English survey: “You will first be presented with two warm-up videos, so that you can understand how the survey works. These will not influence your results. You will get feedback on the warm-up videos, but not on the remaining videos in the survey.”

Instructions presented after warm-up videos:
Norwegian survey: «Oppvarmingen er nå ferdig, og de neste videoene utgjør derfor selve undersøkelsen. Husk at du skal se hver video kun én gang. Du vil fortløpende få spørsmål mellom videoene.»

English survey: “The warm-up is now complete, and the next videos therefore make up the actual study. Remember that you are only going to watch each video once. You will continuously get questions between the videos.”
Appendix F
The Autism Quotients (Short form) – Norwegian version

Instructions: «Vennligst velg de svaralternativene som du mener passer deg best. Det er ingen svar som er rette eller gale. Det er dine oppriktige synspunkter og meninger vi er interesserte i.»

Scale points: Helt enig – Litt enig – Litt uenig – Helt uenig

Items
1. Jeg foretrekker å gjøre ting sammen med andre fremfor på egen hånd
2. Jeg synes sosiale situasjoner er enkle
3. Jeg vil heller gå på et bibliotek enn en fest
4. Jeg finner meg selv sterkere trukket mot mennesker fremfor ting
5. Jeg synes det er vanskelig å få nye venner
6. Jeg nyter sosiale anledninger
7. Jeg liker å treffe nye mennesker
8. Jeg foretrekker å gjøre ting på samme måte gang på gang
9. Det gjør meg ikke opprørt dersom min daglige rutine blir forstyrret
10. Jeg liker å gjøre ting spontant
11. Nye situasjoner gjør meg engstelig
12. Jeg blir ofte sterkt oppslukt av én ting
13. Jeg kan lett holde følge med flere ulike personers samtaler
14. Jeg synes det er lett å gjøre mer enn én ting om gangen
15. Hvis det skjer en avbrytelse, kan jeg bytte tilbake svært raskt
16. Når jeg prøver å forstille meg noe, synes jeg det er lett å lage et bilde i hodet mitt
17. Når jeg leser en historie, kan jeg lett forestille meg hvordan karakterene kan se ut
18. Jeg synes det er lett å dikte opp historier
19. Når jeg leser en historie, synes jeg det er vanskelig å finne ut karakterenes intensjoner
20. Jeg synes det er lett å finne ut hva noen tenker eller føler
21. Jeg synes det er vanskelig å forstille meg hvordan det ville vært å være noen andre
22. Jeg synes det er vanskelig å finne ut folks intensjoner
23. Jeg synes det er lett å leke leker med barn som involverer å late som
24. Jeg legger vanligvis merke til bilskilt eller lignende kjeder med informasjon
25. Jeg er fascinert av datoer
26. Jeg er fascinert av tall
27. Jeg legger merke til mønstre i ting hele tiden
28. Jeg liker å samle informasjon om kategorier av ting

Scoring Procedure:
Responses were scored: (1) Definitely disagree (2) Slightly disagree (3) Slightly agree (4) Definitely agree on the following items: 3, 5, 8, 11, 12, 19, 21, 22, 24, 25, 26, 27, 28.

Responses were scored: (1) Definitely agree (2) Slightly agree (3) Slightly disagree (4) Definitely disagree on the following items: 1, 2, 4, 6, 7, 9, 10, 13, 14, 15, 16, 17, 18, 20, 23.
Appendix G

Autism Quotient (Short form) – English version

Instructions: “Please indicate to what extent the following statements applies to you. There are no right or wrong answers. It is your candid views and opinions we are interested in.”

Scale points: Definitely agree – Slightly agree – Slightly disagree – Definitely disagree

Items

1. I prefer to do things with others rather than on my own
2. I find social situations easy
3. I would rather go to a library than to a party
4. I find myself drawn more strongly to people than to things
5. I find it hard to make new friends
6. I enjoy social occasions
7. I enjoy meeting new people
8. I prefer to do things the same way over and over again
9. It does not upset me if my daily routine is disturbed
10. I enjoy doing things spontaneously
11. New situations make me anxious
12. I frequently get strongly absorbed in one thing
13. I can easily keep track of several different people’s conversations
14. I find it easy to do more than one thing at once
15. If there is an interruption, I can switch back very quickly
16. Trying to imagine something, I find it easy to create a picture in my mind
17. Reading a story, I can easily imagine what the characters might look like
18. I find making up stories easy
19. Reading a story, I find it difficult to work out the character’s intentions
20. I find it easy to work out what someone is thinking or feeling
21. I find it difficult to imagine what it would be like to be someone else
22. I find it difficult to work out people’s intentions
23. I find it easy to play games with children that involve pretending
24. I usually notice car number plates or similar strings of information
25. I am fascinated by dates
26. I am fascinated by numbers
27. I notice patterns in things all the time
28. I like to collect information about categories of things

Scoring procedure:

Responses were scored (1) Definitely disagree (2) Slightly disagree (3) Slightly agree (4) Definitely agree on the following items: 3, 5, 8, 11, 12, 19, 21, 22, 24, 25, 26, 27, 28.

Responses were scored (1) Definitely agree (2) Slightly disagree (3) Slightly disagree (4) Definitely disagree on the following items: 1, 2, 4, 6, 7, 9, 10, 13, 14, 15, 16, 17, 18, 20, 23.
Appendix H
CAPE 42 Positive Subscale – Norwegian version

Instructions: «Disse spørsmålene er skapt for å måle særskilte følelser, forestillinger, og mentale opplevelser. Vi tror at disse er mye vanligere enn det som var tidligere antatt, og at de fleste mennesker har hatt slike opplevelser i løpet av sitt liv. Det er ingen svar som er rette eller gale. Det er dine oppriktige synspunkter og meninger vi er interesserte i. Tenk ikke for lenge på hvert spørsmål - er du i tvil, er ofte den første tanken det beste svaret.»

Scale points: Aldri – Iblant – Ofte – Nesten Alltid

Items

1. Har du noen gang følelsen av at folk kommer med hint om deg, eller sier ting med dobbel betydning?
2. Har du noen gang følelsen av at noe i et blad eller på TV er beregnet spesielt på deg?
3. Har du noen gang følelsen av at enkelte personer er noe annet enn de gir seg ut for?
4. Føler du deg noen gang forfulgt på en eller annen måte?
5. Tror du på at folk har blitt bortført av romvesener, eller at dette kan ha hendt deg?
6. Har du noen gang følelsen av at en konspirasjon er rettet mot deg?
7. Har du noen gang følelsen av at du er forutbestemt til å være en veldig viktig person?
8. Har du noen gang følelsen av at du er en veldig spesiell eller uvanlig person?
9. Hender det at du tror folk kan kommunisere telepatisk?
10. Har du noen gang følelsen av at elektriske apparater, for eksempel datamaskiner, kan påvirke tankene dine?
11. Tror du på hekseri, voodoo eller overnaturlige ting?
12. Har du noen gang følelsen av at du er en berømt, historisk personlighet?
13. Har du noen gang følelsen av at folk ser rart på deg på grunn av utseendet ditt?
14. Har du noen gang følelsen av at tankene dine blir tatt fra deg?
15. Føler du noen gang at tankene i hodet ditt ikke er dine egne?
16. Har tankene dine noen gang vært så livaktige at du var redd andre skulle høre dem?
17. Hører du noen gang dine egne tanker, som et ekko?
18. Har du noen gang følelsen av å være under kontroll av en kraft eller makt utenom deg selv?
19. Hender det at du hører stemmer når du er alene?
20. Hender det at du hører stemmer som snakker sammen når du er alene?
21. Har du noen gang følelsen av at et familiemedlem, en venn, eller en bekjent har blitt erstattet av en dobbeltgjenger?
22. Har du noen som opplevd et mental kollaps, der du blir en annen person?
23. Hender det at du ser gjenstader, folk eller dyr som andre ikke kan se?

Note: All items were scores (1) Aldri (2) Iblant (3) Ofte (4) Nesten Alltid, except items number five, twelve and 22. These are control items, and were not scored.
Appendix I
CAPE 42 Positive Subscale – English version

Instructions: “The purpose of this survey is to assess unique feelings, imaginations, and mental experiences. We believe that these experiences are much more common than previously assumed, and that most people have had such experiences throughout their lives. There are no right or wrong answers. It is your candid views and opinions we’re interested in. Think not too long on each question - if in doubt, it is often the first impulse that is the best answer.”

Scale points: Never – Sometimes – Often – Nearly Always

Items

1. Do you ever feel as if people seem to drop hints about you or say things with a double meaning?
2. Do you ever feel as if things in magazines or on TV were written especially for you?
3. Do you ever feel as if some people are not what they seem to be?
4. Do you ever feel as if you are being persecuted in some way?
5. Do you believe in kidnapping done by aliens, or that you have been kidnapped by an alien?
6. Do you ever feel as if there is a conspiracy against you?
7. Do you ever feel as if you are destined to be someone very important?
8. Do you ever feel that you are a very special or unusual person?
9. Do you ever think that people can communicate telepathically?
10. Do you ever feel as if electrical devices such as computers can influence the way you think?
11. Do you believe in the power of witchcraft, voodoo or the occult?
12. Do you ever feel that you have a famous historic personality?
13. Do you ever feel that people look at you oddly because of your appearance?
14. Do you ever feel as if the thoughts in your head are being taken away from you?
15. Do you ever feel as if the thoughts in your head are not your own?
16. Have your thoughts been so vivid that you were worried other people would hear them?
17. Do you ever hear your own thoughts being echoed back to you?
18. Do you ever feel as if you are under the control of some force or power other than yourself?
19. Do you ever hear voices when you are alone?
20. Do you ever hear voices talking to each other when you are alone?
21. Do you ever feel as if a double has taken the place of a family member, friend or acquaintance?
22. Have you ever experienced a mental collapse where you become another person?
23. Do you ever see objects, people or animals that other people cannot see?

Note: All items were scores (1) Never (2) Sometimes (3) Often (4) Nearly Always, except items number five, twelve and 22. These are control items, and were not scored.
Appendix J

Individual Differences in Anthropomorphism Questionnaire – Norwegian version

Instructions: «Vennligst velg de svaralternativene som du mener er mest korrekte. Det er ingen svar som er rette eller gale. Det er dine oppriktige synspunkter og meninger vi er interesserte i.»

Scale points: 0 (ikke i det hele tatt) – 1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 10 (Svært mye)

Items

1. I hvilken grad har teknologi – instrumenter og maskiner for produksjon, underholdning og produktive prosesser (for eksempel biler, datamaskiner, fjernsynsapparater) - intensjoner?
2. Til hvilken grad har den gjennomsnittlige fisken fri vilje?
3. Til hvilken grad har det gjennomsnittlige fjellet fri vilje?
4. Til hvilken grad opplever et fjernsynsapparat emosjoner?
5. Til hvilken grad har den gjennomsnittlige roboten bevissthet?
6. Til hvilken grad har kyr intensjoner?
7. Til hvilken grad har en bil fri vilje?
8. Til hvilken grad har havet bevissthet?
9. Til hvilken grad har den gjennomsnittlige datamaskinen et eget sinn?
10. Til hvilken grad opplever en gepard emosjoner?
11. Til hvilken grad har det gjennomsnittlige insektet et eget sinn?
12. Til hvilken grad opplever miljøet emosjoner?
13. Til hvilken grad har et tre et eget sinn?
14. Til hvilken grad har vinden intensjoner?
15. Til hvilken grad har det gjennomsnittlige krypsyret bevissthet?

Note: Scored according to scale points. Only responses 0 and 10 had descriptions.
Appendix K

Individual Differences in Anthropomorphism Questionnaire – English version

Information text: “Please indicate to what extent you believe these statements to be correct. There are no right or wrong answers. It is your candid views and opinions we are interested in.”

Scale points: 0 (ikke i det hele tatt) – 1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 10 (Svært mye)

Items

1. To what extent does technology—devices and machines for manufacturing, entertainment, and productive processes (e.g., cars, computers, television sets)—have intentions?
2. To what extent does the average fish have free will?
3. To what extent does the average mountain have free will?
4. To what extent does a television set experience emotions?
5. To what extent does the average robot have consciousness?
6. To what extent do cows have intentions?
7. To what extent does a car have free will?
8. To what extent does the ocean have consciousness?
9. To what extent does the average computer have a mind of its own?
10. To what extent does a cheetah experience emotions?
11. To what extent does the environment experience emotions?
12. To what extent does the average insect have a mind of its own?
13. To what extent does a tree have a mind of its own?
14. To what extent does the wind have intentions?
15. To what extent does the average reptile have consciousness?

Note: Scored according to scale points. Only responses 0 and 10 had descriptions.