

Who Tweets for the autistic community? A natural language processing–driven investigation

Autism
2025, Vol. 29(7) 1740–1753
© The Author(s) 2025



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/13623613251325934
journals.sagepub.com/home/aut



Canfer Akbulut¹ and Geoffrey Bird^{1,2}

Abstract

The formation of autism advocacy organisations led by family members of autistic individuals led to intense criticism from some parts of the autistic community. In response to what was perceived as a misrepresentation of their interests, autistic individuals formed autistic self-advocacy groups, adopting the philosophy that autism advocacy should be led ‘by’ autistic people ‘for’ autistic people. However, recent claims that self-advocacy organisations represent only a narrow subset of the autistic community have prompted renewed debate surrounding the role of organisations in autism advocacy. While many individuals and groups have outlined their views, the debate has yet to be studied through computational means. In this study, we apply machine learning and natural language processing techniques to a large-scale collection of Tweets from organisations and individuals in autism advocacy. We conduct a specification curve analysis on the similarity of language across organisations and individuals, and find evidence to support claims of partial representation relevant to both self-advocacy groups and organisations led by non-autistic people. In introducing a novel approach to studying the long-standing conflict between different groups in the autism advocacy community, we hope to provide both organisations and individuals with new tools to help ground discussions of representation in empirical insight.

Lay Abstract

Some autism advocacy organisations are run by family members of autistic people, and claim to speak on behalf of autistic people. These organisations have been criticised by autistic people, who feel like autism charities do not adequately represent their true interests. In response to these organisations, autistic people have come together to form autistic self-advocacy organisations, or groups in which activists can spread awareness of autism from an autistic point-of-view. However, some people say that autistic self-advocacy organisations do not sufficiently represent the needs of all autistic people. These tensions between organisations and individuals have made it difficult to determine which organisations can make the claim that they represent all autism advocates individuals equally, instead of showing preference to a sub-group within the autism community. In this study, we try to approach this issue using computational tools to see if, in their Twitter posts, both kinds of organisations show a preference for the interests of autistic people or parents of autistic children. We do so by comparing a large body of Tweets by organisations to Tweets by autistic people and parents of autistic children. We find that both kinds of organisations match the interests of one group of autism advocates better than the other. The insight we provide has the potential to inspire new conversations and solutions to a long-standing conflict in autism advocacy.

Keywords

advocacy, autism, natural language processing, Twitter

Introduction

The initial period of autism advocacy was heavily influenced by autism organisations founded by families of autistic children, which have been described as being managed without the input of autistic individuals (McCoy et al., 2020; from this point onwards, organisations that are

¹University of Oxford, UK

²University College London, UK

Corresponding author:

Geoffrey Bird, Department of Experimental Psychology, University of Oxford, Anna Watts Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, United Kingdom.

Email: geoff.bird@psy.ox.ac.uk

not self-designated as led by autistic people will be referred to as non-self-advocacy groups, or nASA for short).¹ In the early days of their existence, nASA organisations often adopted a biomedical view of autism, conceptualising it as an illness that – with adequate research – could be treated to lessen the severity of its symptoms, or cured entirely (Krcsek, 2013). Communications from these were often targeted at family members of autistic children to supply parents with resources, information and support as they navigated their children’s diagnosis. In public statements, prominent charities have described autism as a ‘national emergency’ and a ‘monumental health crisis’ (Wilson, 2013), characterising it as a condition that greatly obstructs the ‘hopes’ and ‘dreams’ of those afflicted and causes severe ‘struggle’, ‘embarrassment’ and ‘pain’ for the families involved (Autistic Self Advocacy Network (ASAN), 2009).

nASA groups have drawn criticism from autistic self-advocates, who object to what they have described as nASA groups’ stigmatising depictions of autism, as well as their failure to include autistic people in decision-making processes within their organisations (ASAN, 2016). Autistic self-advocacy (ASA) organisations emerged in the 1990s, challenging the dominance autism charities had over popular narratives of autism (Ortega, 2009). Drawing from the broader disability advocacy movement (Ne’eman & Bascom, 2020), autism self-advocates adopted the social model of disability rather than the biomedical construction of autism (Krcsek, 2013). Instead of characterising autism as inherently pathological and thus requiring a cure, the social model promotes neurodiversity, or a desire for autistic people to be included into society without changing aspects of their cognition or behaviour to be accepted (Kapp et al., 2013; Pellicano & den Houting, 2022). ASA organisations maintain that autistic people, and not their neurotypical family members, should be at the forefront of decisions on research, policy and public communication surrounding autism (Bertilsdotter Rosqvist et al., 2015; Ne’eman & Bascom, 2020).

The popularity of ASA groups and the views they express prompted a debate around the question of which groups speak for autistic individuals. Since their inception, ASA organisations have asserted that nASA groups make wrongful claims of representativeness, as the views of autistic people were scarcely considered in their initiatives (Chapman & Veit, 2020; Fellowes, 2020; Rosenblatt, 2018). However, family members of autistic children have raised doubts about ASA organisations’ claims of representativeness (Richardson & Sharp, 2020), arguing that ASA groups, as spokespeople for the entirety of the autistic community, have a responsibility to represent a diverse – and occasionally conflicting – set of interests (McCoy et al., 2020).

Parents of autistic children with more severe intellectual, social and behavioural impairments have reported feeling alienated by ASA organisations (Haney, 2018), in large part due to their vocal opposition to policies aimed at addressing parents’ supervision and safety concerns. For example, a proposal to subsidise personal locating devices for families with autistic children prompted an outcry from leading ASA organisations, amid parental insistence that such devices could prevent children from getting lost or wandering into dangerous situations (Joseph, 2011). In prioritising autonomy and independence outcomes for autistic adults, parents maintain, ASA organisations neglect to consider the unique needs and concerns of autistic people unable to speak for themselves (Richman, 2020).

In an exploration of what it means for an organisation to ethically advocate for the autistic community, McCoy and his colleagues (2020) suggested that ASA and nASA organisations – as parties assuming a representative role not through electoral means, but through self-appointment – engage in practices of ‘partial representation’. For example, while ASA groups claim to speak for all autistic people, they may be preferentially amplifying the voices of autistic people who are capable of speaking for themselves. To ensure that they do not overlook autistic people with greater support needs in their advocacy, McCoy et al. recommended that organisations seek endorsement from parents and caregivers with ‘firsthand knowledge of the needs of autistic people with severe disabilities’.

In response to this recommendation, some argued that parents could not reliably represent their autistic children’s interests (Benjamin et al., 2020), suggesting that, for organisations, representational priorities should lie with amplifying the expressed views of autistic people. Others argued that parents deserve to be represented in their own right, not solely as proxies for their children (Buturovic, 2020), affirming the role of the parent as a stakeholder in autism advocacy. It is clear, then, that the community has yet to reach a consensus on what constitutes ethical, fully representative autism advocacy, leaving a legacy of contradictory arguments for organisations to follow. The question remains: who speaks for the autistic community?

This study introduces a new approach to investigating this critical question. By analysing social media data with natural language processing (NLP) tools, we compare the online expressions of autistic self-advocates and parents of autistic children with messages from organisations that engage in self-appointed representation practices: nASA groups and ASA organisations. While this method cannot definitively prove nor disprove claims of partial representation, the many comparisons drawn between language used by individuals and organisations may capture a snapshot of representational patterns in online advocacy. This method allows us to explore criticisms against both ASA groups and nASA organisations, providing empirical evidence to

support or challenge these claims. The empirical methods introduced are not infallible: social media data may not fully reflect the diversity of the autism advocacy community, and language modelling results introduce ambiguities in interpretation. Even so, we believe the study offers a valuable methodological contribution to the field, providing a new way to explore the complex dynamics of representation within the autism community.

Autism advocacy on Twitter

The focus of this work is on online autism advocacy, and in particular, advocacy on the platform Twitter (now known as X). The advent of Internet-based communication created widely accessible channels for disability advocacy, enabling users to bond over shared experiences and rally against the stigmatisation, prejudice and exclusion they face in a largely able-bodied and neurotypical society (Bowker, 2008; Bowker & Tuffin, 2004; Jaeger, 2022). Online spaces dedicated to the discussion of autism by autistic people allowed individuals to forge a positive autistic identity centred around the celebration of their neurodiversity. In Internet forums, autistic people embraced their idiosyncracies rather than framing them as symptoms that must be remedied; asserted agency over their way of life in a rejection of the viewpoint that they are victims of their condition; and developed new terminology to distance themselves from language that positions them as lesser than their non-autistic peers (Parsloe, 2015).

Although autism advocacy is represented across many social media platforms, Twitter was selected for this study for three reasons: (1) at the time the study was conducted, it offered generous API (Application Programming Interface) access for academic researchers, allowing us to efficiently collect large-scale datasets; (2) its API supported features such as keyword-based searches of users and Tweets that were critical to the design of the study; and – most critically – (3) Twitter has long been a critical platform for online autism advocacy. As online advocacy spread from lesser-known forum boards to mainstream social media networks, Twitter emerged as a space for autistic individuals to make themselves heard by other autistic people, non-autistic people and autism organisations alike (Trevisan, 2016). Using the #ActuallyAutistic hashtag, Twitter users publicly identify themselves as autistic, sharing personal experiences with the intention of countering overly medicalised narratives that dominate public perceptions of autism ([@autisticadvocacy], 2014). Critically, Twitter is a powerful tool for dissent and protest, providing autistic individuals with a platform to express collective opposition to social and medical policies that they consider harmful to autistic people. In recent years, autistic individuals have popularised the #SayNoToABA hashtag to highlight the underexplored adverse effects of

applied behavioural analysis (ABA), a controversial autism treatment ([@AutSciPerson], 2021). Similarly, the #StopTheShock hashtag gained prominence in response to reports of an institutional care centre routinely administering electric shocks to its autistic in-patients as a method of behavioural control (ASAN, 2022). This study leverages ‘Tweets’ (short public messages on the Twitter platform) by autistic users to capture overall patterns in their language use, which may be indicative of their public values and beliefs.

Much like autistic individuals, parents of autistic children rely on Internet-based communication in their research into their children’s diagnoses. Parents may turn to other parents of autistic children on the Internet for educational resources and emotional reassurance, building a community based on their shared experiences of parenting autistic children (Fleischmann, 2005; Reinke & Solheim, 2015) and forging new identities as ‘autism parents’ (Hall et al., 2016). On Twitter, parents have established hashtags such as #AutismMom, #AutismDad and #AutismParent to document their experiences, promote their interests and make their needs known.

Corporate social responsibility

Organisations are also users of Twitter. Organisations seeking to communicate their commitment to conducting business in an ethical manner commonly do so on social media platforms, allowing for highly visible displays of corporate social responsibility (CSR). Companies that broadcast CSR-related messaging do so to cultivate the belief among stakeholders that their engagement with CSR is authentic (Hur et al., 2020; Lock & Seele, 2017). In a comprehensive study exploring the virtual avenues for CSR-related messaging, it was found that a majority of corporations use Twitter to keep stake-holders informed of their CSR activities (J. Yang et al., 2020). Within companies in the non-profit sector – an umbrella term under which many registered ASA and nASA organisations fall – negative perceptions of CSR have been found to decrease the perceived trustworthiness of the organisation (Lin-Hi et al., 2015). This erosion of trust may have significant ramifications for stake-holders’ willingness to support the organisation’s activities, an effect that is particularly detrimental for organisations that are financially reliant on their stake-holders’ goodwill. A company using social media to appeal to an audience may benefit from demonstrating that its social responsibilities are aligned with those that it positions as stake-holders (Connelly et al., 2011); establishing a ‘good fit’ with key actors may be especially critical for organisations whose mission is centred around the promotion of an altruistic social cause (Graff Zivin & Small, 2005; Jamali & Mirshak, 2007), such as autism awareness and acceptance.

Applying the lens of CSR requires the identification of a body of stake-holders, to whom the organisation tries to appeal to through its social media activity. However, identifying the primary stake-holders of autism advocacy organisations is not a trivial task. Some organisations may describe their constituents as a narrower group than all people involved in autism advocacy. For example, in their page on ‘What We Believe’, the Autistic Self-Advocacy Network states ‘that autistic people need to be involved whenever autism is discussed’, centering autistic self-advocates in their mission. Other organisations may make implicit claims of inclusiveness: the Autism Speaks home page features resources for both autistic people and parents of autistic children side-by-side, which could be interpreted as an intention to equally represent and service both interest groups.

We do not have insight into an organisation’s internal decision-making processes beyond materials they choose to externalise, so we cannot make definitive claims of who organisations intend to represent. Considering claims of biased advocacy and exclusionary attitudes have been levied against both ASA and nASA organisations; however, it may be argued that some hold an expectation that both kinds of organisations should aim to represent a broad range of stake-holders, encompassing many groups implicated in autism advocacy, research and policy. In our methodology, we compare organisational Tweets to Tweets by both groups of individuals to capture potential displays of preferential alignment, whether this alignment is intentional or not.

Drawing on foundational theories from the organisational CSR literature (Graff Zivin & Small, 2005; Jamali & Mirshak, 2007), we assume that ASA and nASA groups alike are motivated to display alignment with whichever group they consider their core audience members. Accordingly, if nASA groups deem parents of autistic children to be their most important audience members, their values – as reflected by the language used in their Tweets – should overlap more significantly with those of parents than autistic individuals. Likewise, if ASA groups are primarily interested in appealing to autistic Twitter users, their expressed values would likely register as more similar to the beliefs of autistic individuals than to parents of autistic children.

Informed by the direction in which criticisms of partial representation have been raised, two specific predictions are made: (1) the online expressions of ASA organisations will be more similar to expressions of autistic self-advocates than to those of parents of autistic children and (2) the online expressions of nASA organisations will be more similar to the expressions of parents of autistic children than to those of autistic individuals. Because the chosen NLP method represents semantic patterns in language use (Goldberg & Levy, 2014), rather than syntactic or frequency-based approaches, we posit that groups whose

language use is ‘more similar’ make expressions that are ‘closer in meaning’, though we are unable to collect additional evidence – such as self-report data – to support this claim.

Methods

We gauged the alignment of an organisation’s public messaging with different groups by analysing how similarly they use language in their Tweets. Specifically, we derived a measure of similarity between the semantic contexts in which shared words occur across the Tweets of organisations and individuals. We produced measures of semantic similarity in three steps: (1) constructing a corpus of Tweets for each organisational and individual group of interest; (2) generating word embeddings as a means of calculating semantic similarity between groups; and (3) conducting a specification curve analysis (SCA) as a robustness check of the hypothesis testing procedure. For a visual summary of methods, see Figure 1.

Constructing corpora

For comparisons across organisational and individual online expressions to be made, four distinct corpora of Tweets were obtained, each authored by one of the four groups of interest. Because no such corpora are readily available, they were constructed for the purposes of the study. All data were collected between March and August of 2022.

First, it was necessary to establish a dictionary of words and phrases used in reference to autism, both by autistic people engaging in self-description and by non-autistic people describing autistic others. A list of autism-relevant keywords was compiled through an informal survey of graduate students at an autism-centred research group (who were both autistic and non-autistic). For the full list of keywords and more detail on the selection process, see Supplementary Materials A.1.1.

To construct a corpus of Tweets by organisations, autism-relevant keywords were applied to the Twitter API’s ‘user search’ function, a relevance-based search that returns a 1000 public user accounts per keyword. Twitter accounts were manually labelled as ASA groups if their biographies (1) did not contain references to the account being owned by an individual (e.g. account name or biography contains personal details) and (2) contained references to self-advocacy. If a Twitter account did not belong to an individual but did not mention self-advocacy, it was labelled as a nASA group. Twitter accounts belonging to users identified as individuals were discarded. We note that although this process introduces the possibility of misclassification – that ASAs are labelled as nASAs – we chose to take a conservative approach with identifying self-advocacy groups, as they generally adhere to more

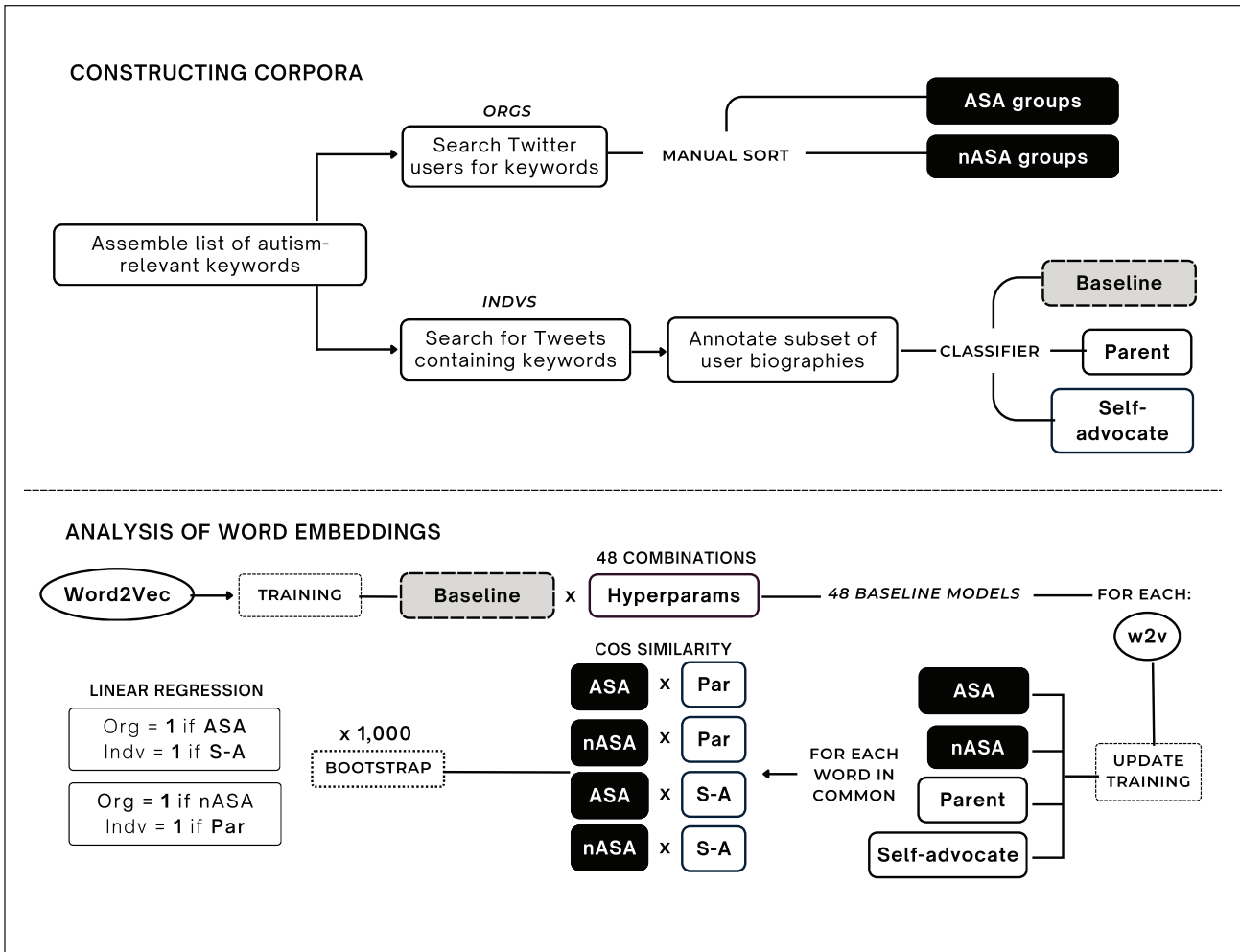


Figure 1. Summary of procedure for finding semantic similarity between common words in Tweets across autism-relevant organisational and individual groups. ‘w2v’ refers to a Word2Vec model.

specific criteria on who can assume a representative role over the autistic community (Ne’eman & Bascom, 2020; Zuber & Webber, 2019). We also acknowledge that our approach of distinguishing between groups and individuals may increase the possibility of false positives: that non-groups are labelled as groups. We choose this approach because, at the time the data were collected, Twitter API limit restricted researchers to collecting up to 3200 of a users’ most recent Tweets, so we employ a more permissive selection approach for groups to ensure that we collect a sufficiently large sample of Tweets on which to fine-tune word embedding models. These Tweets were then sorted using the manual labels provided to create two distinct corpora: Tweets by ASA and nASA groups. The public Twitter handles of accounts that constituted our organisational set are available upon request, which is a measure we take to preserve the privacy of users.

While it could be presumed that the majority of Tweets published by autism-centred groups would pertain to topics relevant to autism, the same assumption cannot be

extended to autistic individuals and parents of autistic children. Consequently, the procedure for constructing Tweet corpora differed for individuals. First, a large number of Tweets containing autism-relevant keywords were collected, excluding re-tweets (re-posts of another user’s materials) and Tweets not authored in the English language. The biographies of the Tweets’ authors were retrieved. A random subset of 20,165 user biographies were labelled by a human annotator, such that each biography was deemed to belong to a user who self-identified as autistic, a parent of an autistic child or neither. Twitter users already represented in corpora of organisational Tweets were excluded from this sample.

Next, a number of supervised classification models were trained on the labelled examples to identify which category a user belonged to based on their biography (for more detail on the annotation and model selection process, see Supplementary Materials A.1.2 and A.1.2.1). The model with the highest test accuracy (the highest proportion of correct labels assigned to biographies that it had not

been trained on) was used to classify the rest of the user biographies. The corpus designation of each Tweet's author – autistic, parent of autistic child or non-autistic – was therefore determined by a human rater, or by a model trained on human ratings.

At the end of the corpus construction process, discrete collections of Tweets were available from the four groups of interest: ASA groups, nASA groups, autistic self-individuals and parents of autistic children. Tweets that were retrieved during the keyword-based Tweet search, whose authors did not belong to any these groups, were stored in a separate 'baseline' corpus for later use in pre-training the language model.

Word embeddings

Advances in the field of NLP have produced sensitive methods of generating numerical representations of words to capture their underlying semantic qualities. Within these methods, a word's 'embedding' – a vector representation that reflects the many contexts in which it occurs – can be thought of as its coordinates in a hyperdimensional semantic space, where each dimension reflects a semantic quality (Kusner et al., 2015). Words that have a tendency to occur near each other are far likelier to have a stronger semantic relationship than words that rarely inhabit the same context; this property of language can be envisioned as the physical distance between two embeddings, such that the closer together two embeddings are situated in a semantic space, the more shared meaning exists between the words.

A word's embedding is not fixed and unchanging: a language model determines a word's embedding by analysing where it occurs 'within the corpus on which it is trained'. This dynamic property of language models has been leveraged by social scientists to capture sociocultural attitudes and trends through language use. Findings that a language model trained on mainstream news sources produces associations between occupations and gendered words along stereotypically gendered lines, such that 'woman' relates more closely to 'secretary' than 'engineer', have shed light on a collective culture of gender bias (Caliskan et al., 2017). A powerful example of a language model's ability to parallel the attitudes, sentiments and concerns of its training documents was also presented in the study by Garg et al. (2018), in which embeddings trained on 19th-century American literature demonstrated an increased association between Asian-related terms and derogatory adjectives following waves of Asian immigration into the United States.

A language model's unique ability to capture the meaning of words within a corpus has also been used to compare attitudinal similarities and disparities between groups. Lawson et al. (2022) illustrated the effect that hiring female chief executive officers (CEOs) can have on gender stereotypes in organisational language, such that positive

associations between words pertaining to femininity and agency emerged in organisations that experienced a leadership transition to a female CEO, relative to organisations that did not.

As embedding representations, words that frequently share the same context are positioned closer to one another in semantic space. It follows that if identical words trained on two distinct corpora are situated close to one another, there is a great overlap in the contexts in which these words are used between corpora. Furthermore, greater similarity in the positions of an identical word between Corpus A and Corpus B, as compared to Corpus A and Corpus C, suggests that there is greater shared meaning in the way the word is used in Corpora A and B. If, for example, the word 'autism' frequently occurs in a biomedical context in the Tweets of nASA groups and parents of autistic children, but not in the Tweets of autistic individuals, this would be reflected in the greater distance between embeddings of 'autism' in nASA and self-advocate, relative to nASA and parent, comparisons. Thus, disparities in organisational representativeness could be illustrated through organisation/individual comparisons for each identical word across the four corpora of interest.

Estimates of shared meaning between autism-relevant words in comparisons across organisational and individual corpora were recovered in the following steps. Prior to training, a text-cleaning procedure was applied to all Tweets (more details in Supplementary Materials A.1.3). Baseline embeddings for many words were generated using the gensim package on Python (Rehurek & Sojka, 2011) to train an unsupervised NLP algorithm, Word2Vec (Mikolov et al., 2013), on the baseline Tweet corpus. Word2Vec produced an equal-length multi-dimensional vector for unique words in the baseline corpus. Next, the baseline model was trained on each of the corpora of interest separately, yielding four distinct sets of word embeddings. In this step, the pre-trained language model will adjust the semantic placement of words from the baseline corpus depending on how the word is used in the new training corpus; words used in similar contexts across two corpora are adjusted in similar directions, resulting in smaller distances between embeddings.

Embeddings of words not present across all corpora were removed, and the distances between embeddings of identical words between organisations and individuals (nASA groups × parents of autistic children, nASA groups × autistic individuals, ASAG groups × parents of autistic children, ASAG groups × autistic individuals) were computed. Cosine similarity, or the cosine of the angle formed by two numerical vectors, served as the distance metric and was calculated between each identical word embedding pair using the SciPy package on Python (Virtanen et al., 2020). Cosine similarity can assume any value between -1 and 1, such that a cosine similarity of -1, 0 or 1 represents an opposite, orthogonal or identical relationship between two vectors. Note that, while negative cosine similarities indicate that two word embeddings are semantically

Table 1. Hyperparameter space for SCA.

Hyperparameter name	Input values
Word vector dimensionality	100, 300, 500
Sampling window	5, 8
Minimum count of words	30, 50
Training algorithm	skip-gram continuous bag-of-words
Output layer function	hierarchical softmax, negative sampling ^a

SCA: specification curve analysis.

^aNegative sampling value set to 10 noise words, as per gensim recommendations.

dissimilar, values close to zero signal a lack of semantic similarity and dissimilarity.

Following this procedure yielded a set of cosine similarities, such that each similarity metric was associated with a single word for a given individual and organisation comparison. Inspired by the analyses in the study by Lawson et al. (2022), a linear regression analysis was applied to test Hypothesis 1 – that the Tweets of ASA groups organisations are closer in meaning to (and therefore have greater, more positive cosine similarity scores with) the Tweets of autistic self-advocates than to those of parents with autistic children – using the equation and binary coding specification in Equation (1).

$$\begin{aligned}
 \text{Cos Sim} &= \beta_1 \text{Org} + \beta_2 \text{Ind} + \beta_3 (\text{Org} * \text{Ind}) \\
 \text{Org} &= \begin{cases} 1, \text{ if ASA} \\ 0, \text{ if nASA} \end{cases} \\
 \text{Ind} &= \begin{cases} 1, \text{ if Autistic Person} \\ 0, \text{ if Parent of Autistic Child} \end{cases}
 \end{aligned} \tag{1}$$

Of particular interest is the β_3 interaction coefficient, which represents the extent to which words in the Tweets of ASA organisations demonstrated a greater overlap in shared meaning with words used by autistic individuals, as compared to those used by parents of autistic children.

To test the second prediction – that the Tweets of nASA groups are closer in meaning to the Tweets of parents of autistic children than to those of autistic individuals – the binary coding specifications are reversed (Equation (2)). The β_3 coefficient now represents the degree to which words in Tweets by nASA groups are used in more similar contexts to words in Tweets by parents of autistic children relative to Tweets by autistic individuals.

$$\begin{aligned}
 \text{Cos Sim} &= \beta_1 \text{Org} + \beta_2 \text{Ind} + \beta_3 (\text{Org} * \text{Ind}) \\
 \text{Org} &= \begin{cases} 1, \text{ if nASA} \\ 0, \text{ if ASA} \end{cases} \\
 \text{Ind} &= \begin{cases} 1, \text{ if Parent of Autistic Child} \\ 0, \text{ if Autistic Person} \end{cases}
 \end{aligned} \tag{2}$$

Specification curve analysis

We note that a word's embedding depends not only on the training corpus, but also on the hyperparameters that are selected for the language model. A large-scale study of Word2Vec language model performance found that adjustments to a variety of hyperparameters had a pronounced impact on its performance on a recommendation-based task (Caselles-Dupré et al., 2018). Within the context of this study, the possibility that a significant effect is an artefact of hyperparameters of the language model could not be dismissed. An SCA was, therefore, employed to ensure that recovered β_3 parameters were not due to chance.

First, a number of baseline models were trained, varying across multiple dimensions of hyperparameters. The full hyperparameter space is outlined in Table 1, such that each possible combination of hyperparameter specifications was utilised, yielding 48 ($3 \times 2 \times 2 \times 2 \times 2$) baseline models. Each baseline model was updated on the four corpora of interest separately. For each set of updates on the baseline model, only embeddings of words that occur across all four corpora were retained, and cosine similarities of identical words between corpora pairings (nASA \times parent, nASA \times self-advocate, ASA \times parent, ASA \times self-advocate) were produced, yielding four cosine similarity scores per shared word per baseline model.

A 1000 bootstrapped samples were drawn from these datasets of cosine similarities. For each iteration, 10,000 words were sampled and their z-standardised cosine similarities across organisation and individual pairings used to produce two β_3 parameters, indicating the similarity of ASA groups and autistic individuals (Equation (1)) and nASA groups and parents of autistic children (Equation (2)).

A total of 48,000 (48 specifications \times 1000 samples) pairs of β_3 estimates were recovered, then plotted by effect size. As a simple benchmark, the cosine similarities of each bootstrapped sample were randomly shuffled so that the organisation and individual labels no longer corresponded to the correct similarity score. The 'random' β_3 estimates generated from the shuffled datasets were plotted alongside the specification curves as a point of comparison.

Statement on community involvement

While there was little community involvement in the design of the study, the authors make every effort to include the viewpoints of autistic self-advocates in the data collection and analyses processes. We do so by ensuring our search word approach for identifying autistic self-advocates on Twitter includes language autistic self-advocates – rather than autism researchers – are likely to use (Supplementary Materials A.2.1). We also devote a significant amount of annotation time (Supplementary Materials A.2.2) to capturing as many self-identifying autistic Twitter users in our scraping dataset as possible to increase

Table 2. Summary statistics of Tweet corpora.

Corpus	Tweets	Users	Words	Unique words	Words: Tweet	Tweets: User
Baseline	5,110,388	1,469,851	64,664,557	516,144	12.65 (5.83)	3.48 (37.77)
Indv: self-advocate	474,149	10,687	6,754,644	110,628	5.44 (2.96)	44.36 (192.2)
Indv: parent	288,638	10,827	3,426,058	84,239	5.52 (2.97)	26.66 (391.62)
Org: nASA	557,554	137	6,210,788	97,895	5.69 (2.83)	4069.73 (4987.11)
Org: ASA	342,301	134	4,533,153	86,934	5.69 (2.91)	2554.49 (5646.53)

Average of ratios appear as Mean (SD). 'Unique words' represent the number of unique tokens in each corpus after data-cleaning procedures are applied. nASA: non-autistic self-advocacy; ASA: autistic self-advocacy.

the size and diversity of the training, validation and test sets used to develop and evaluate our classification approach.

Results

Pre-registration and code availability

All documents related to the study's pre-registration can be found on the Open Science Framework.² Raw data cannot be shared due to the privacy restrictions imposed by the user agreement for Twitter API Academic Research Access, but cosine similarity scores are made available. All codes can be found on GitHub.³

Sample

Using Tweets from organisations and individuals in autism advocacy, the z-standardised cosine similarities of shared words across pairings of organisational and individual corpora were retrieved. That is, for a shared word that occurs across all four corpora, we calculate the distance between the way the word is embedded in corpus-specific models. Calculating this distance tells us how different the semantic context in which the word occurs is between corpora. For instance, if the distance between the embeddings for 'autism' between ASA groups and autistic self-advocates is smaller than the distance for ASA groups and parents of autistic children, we may conclude that 'autism' occurs in more similar semantic contexts in the former pairing than the latter.

The interaction terms from Equations (1) and (2) were computed, indicating the extent to which an organisation had greater semantic similarity with one individual group than the other. As described in the Methods, from our full dataset of all cosine similarity for each shared word, we drew samples of 10,000 words for 1000 attempts, and fitted the samples to Equations (1) and (2). We extracted the β_3 value from both equations, representing the similarity of language use between organisations and individuals. We then plotted interaction coefficients in order of increasing effect size to create a specification curve, which takes into account the potential variability introduced by the hyperparameter space outlined in Table 1. Finally, distributions

of bootstrapped β_3 coefficients were compared against a random benchmark, which was created by shuffling cosine similarity values so that they no longer corresponded to the actual organisation and individual corpus labels. The random benchmark is intended to demonstrate what the results would look like if observed semantic similarity were due to random noise, providing a simple comparison to semantic similarity values belonging to true pairings of organisations and individuals.

The total number of Tweets, users, words, and unique words in each corpus are summarised in Table 2. In addition, corpus-specific average ratios of words per Tweet and Tweets per user are presented.

Specification curve analysis

In Figure 2, interaction coefficients (β_3) – each representing the semantic similarity of a sample of 10,000 words used by organisations and individuals – are plotted in order of effect size. The specification curves for the random benchmark are presented, representing semantic similarity values for words with random organisation and individual corpus labels.

Interaction terms resulting from the ASA groups \times self-advocates binary coding specification (Equation (1)) demonstrate that comparisons of word embeddings between ASA organisations and autistic individuals yielded higher cosine similarity scores than comparisons between ASA groups and parents of autistic children. Simply put, the same words were used in more similar semantic contexts between ASA groups and autistic individuals than ASA groups and parents of autistic children. All β_3 terms along the specification curve were positive (Mean=0.01471, SD=0.00059), such that the 95% confidence interval did not include zero (CI₉₅=[0.01468, 0.01475]).

A similar pattern of findings was highlighted by the specification curve for the nASA groups \times parent coding specification (Equation (2)), which represented the difference in the cosine similarities between nASA groups and parents of autistic children as compared to nASA groups and autistic self-advocates in the true data. All β_3 terms were positive, non-zero values (Mean=0.01471, SD=0.00059, CI₉₅=[0.01468, 0.01475]). These findings show that language was used more similarly between

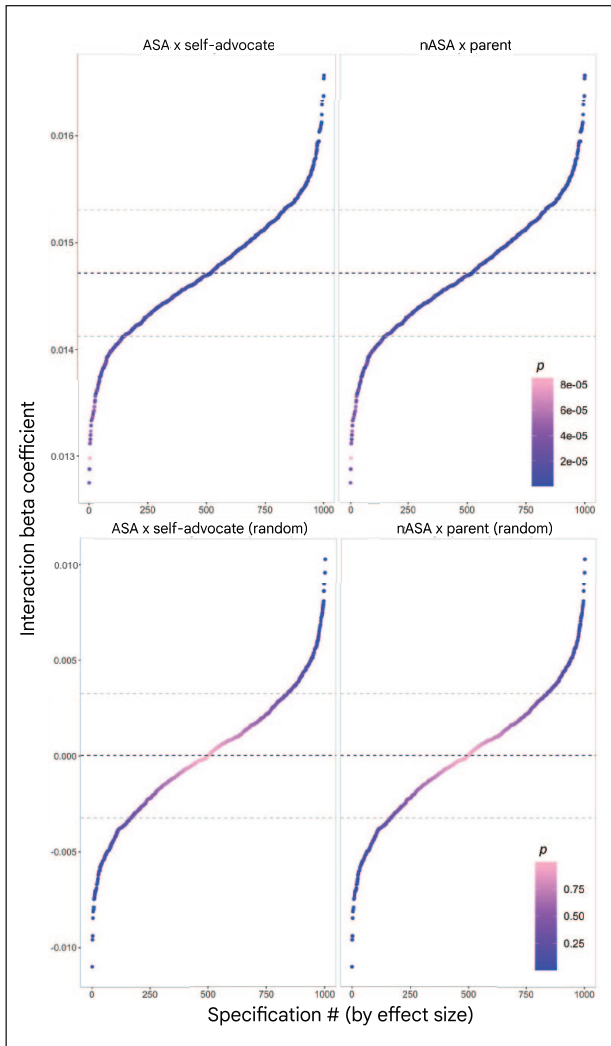


Figure 2. Specification curve of interaction coefficients produced by regressing z-standardised cosine similarity per shared word on binary-coded organisation and individual labels. Below, specification curve of interaction coefficients produced by regressing randomly shuffled, z-standardised cosine similarity per shared word on binary-coded organisation and individual labels is shown. Colour-coded p -values denoted by row. Average of β_3 denoted by blue line; ± 1 SD denoted by light grey lines.

nASA groups and parents of autistic children than nASA groups and autistic self-advocates.

Of all β_3 coefficients produced by randomly shuffled datasets using the ASA \times self-advocate coding specification, only 0.1%, 0.7% and 4.7% were significant at an alpha level of 0.001, 0.01 and 0.05, respectively. The same was true of the β_3 terms yielded by the nASA \times parent specification. In contrast, all p -values corresponding to the true β_3 terms were significant at an alpha level of 0.001.

A two-tailed Welch's t -test indicated a significant difference between the average of the ASA \times self-advocate β_3 distribution and the average of its randomly shuffled counterpart, $t(1064.7)=140.59$, $p < 0.001$, Cohen's $d=6.20$. A significant difference was also found between the average β_3 value of the nASA \times parent curve and the average β_3 of its random benchmark model using a two-tailed Welch's t -test, $t(1064.7)=140.59$, $p < 0.001$, Cohen's $d=6.20$. These findings demonstrate that findings in the top row of Figure 2 are unlikely to be due to random noise, as they are significantly different to chance values, shown in the bottom row.

Discussion

The decades-long conflict between organisations and individuals in autism advocacy has thrust the greater autism acceptance movement into an uncertain future, continuing to impede cooperation within the autism advocacy community and compromising its ability to arrive at mutually beneficial research and policy decisions (Elster & Parsi, 2020; Siegel, 2020; Tabor, 2020). We find that comparing large corpora of Tweets made by organisations and individuals reveals patterns of language similarity consistent with the direction of previous criticisms of organisational partial representation. Adopting the view that word embeddings encode values and attitudes held at the group-level (Caliskan et al., 2017; Garg et al., 2018; Lawson et al., 2022), these findings indicate that ASA groups express more similar views to autistic self-advocates, while nASA groups share values more consistent with those of autistic parents.

Whether or not this study reveals a failure in representation is contingent on many unknown – and, if relying solely on our methods, unknowable – variables. For one, organisations do not always explicitly state the bounds of the representational role in autism advocacy. Without additional insight into organisational decision-making, we cannot conclude if patterns of partial representations are indicative of a failure to meet representational commitments, or a sign of the intentional prioritisation of certain stake-holders over others. We also do not have access to balanced perspectives on the expected level of representation behaviours of autism advocacy groups. If an organisation deliberately and openly prioritises one group over another, is it considered exempt from criticisms of partial representation? On the contrary, if an organisation construes its autism advocacy mission more broadly, is the onus of avoiding partial representation higher than for organisations with more specific inclusion aims? Future studies may build upon the methods introduced in this article by factoring in organisations' own representational claims (e.g. through mission statements and promotional materials) to highlight the difference between purported

and observed representation. They may also incorporate insights on public or group-level attitudes towards expected patterns of organisational representation to add nuance to the implications of our empirical results.

Thus far, contributions to the debate on partial representation were largely based on personal views, requiring many normative assumptions to be made in the process. This work provides a method through which these subjective assertions can be corroborated by real-life data from all groups implicated in the debate. The impact of our methods and findings is twofold: first, we believe the snapshot of autism advocacy on Twitter can help enrich the autism advocacy community's understanding of representation as it occurs on social media platforms. Second, we hope to provide a reproducible method that allows both organisations and individuals in autism advocacy to monitor the overall state of online organisational representation. For organisations, this method may enable oversight over the alignment of their stated missions with observable instances of representation. This empirically grounded oversight, in turn, may catalyse meaningful reflection and paths of recourse to be taken if inconsistencies between representational goals and outcomes are revealed. Similarly, groups of individuals may benefit from these methods as a means of holding self-appointed representational bodies accountable to their stated representational aims. Before recommending the wide adoption of this method, however, we must consider the limitations imposed by the inherently limited representation offered by sampling from social media. We also discuss difficulties with relying on error-prone machine-learning methods to automatically identify individual stakeholders in the partial representation debate.

Limitations of our sampling strategy

The results offer an insight into discussions of representativeness as expressed through the official Twitter communication of autism-centred organisations and the personal Tweets of users in autism advocacy. As predicted, language in Tweets authored by ASA organisations are more closely associated with language used by autistic individuals than parents of autistic children. Also as predicted, the language used by nASA groups was more similar to language Tweeted by parents of autistic children than to that of autistic self-advocates. Taken at face value, both findings lend weight to claims of partial representation levied against both self-advocacy groups and charities. However, in stating the implications of our findings, we must consider how conducting studies based only on data collected from social media may result in a limited and skewed sample of both autistic individuals and parents of autistic children.

An obvious consideration is the potential exclusion of important subgroups of autistic individuals from the study

sample. For those having an issue with ASA groups, one of the most frequently raised points of contention is their purported failure to consider the interests and needs of autistic populations with higher levels of disability (McCoy et al., 2020). Autism – insofar as the term serves as a short-hand for the medical diagnosis of autism spectrum disorder – manifests in a wide range of ‘symptoms, skills, and levels of disability’ (National Institute of Neurological Disorders and Stroke, 2018). Autistic adults may display anywhere from no to severe intellectual impairment (Howlin & Moss, 2012; Lyall et al., 2017), with higher levels of disability resulting in a greater need for external care and support (Sanchack & Thomas, 2016). In their campaigns for increased autonomy, independence and freedom for autistic individuals (Ne’eman, 2011, 2021), ASA groups may have alienated a segment of autism advocates – most notably, parents of children with severe impairments – who hold that self-advocates’ demands are often at odds with legislation, policy and healthcare decisions that would meet the safety and care needs of autistic people with higher levels of disability (Joseph, 2011; Lutz, 2013, 2021). While self-advocates have publicly contested this claim, defending the neurodiversity movement as one that is predicated upon the inclusion of all levels of disability (Des Roches Rosa, 2013; Ne’eman & Bascom, 2020), it remains a highly polarising topic in discussions of organisational representation.

The findings of this study revealed that ASA groups use language in more similar ways to autistic individuals. Given the philosophical underpinnings of self-advocacy, some may argue that self-advocacy organisations justly prioritise autistic over non-autistic interests. Underlying this statement, however, is the assumption that the autistic Twitter users represented in the study, whose Tweets are used to construct the ‘autistic self-advocate’ corpus, accurately reflect the autistic community at large. By virtue of the study design, in which a large volume of Tweets were collected without direct contact between the researcher and Twitter users, it is impossible to determine the extent to which the autistic sample includes autistic individuals with greater levels of disability.

Of adults who are diagnosed with an autism spectrum disorder in childhood, it is estimated that about one-third will experience severe difficulties with communication (Koegel et al., 2020; Rose et al., 2016; Tager-Flusberg & Kasari, 2013), including effective written communication (Dockrell et al., 2012, 2014). Moreover, the lack of accessibility and inclusive design features on social media may feel prohibitive to some autistic people navigating such platforms (Glumbić et al., 2022; Kofmel, 2019). Taken together, these factors highlight several ways in which impairments related to autism may present barriers to engagement with self-advocacy on Twitter, casting doubt on the diversity of the autistic experience that is captured by sampling the Tweets of self-advocates.

Moreover, the views of users who participate in advocacy on social media platforms may deviate substantially from the average view of the wider population from which they are sampled. Previous studies have shown that online activists are likely to display attitudes and behaviours that distinguish them from the average individual. For example, in attempts to characterise voter behaviours by sampling from Twitter, the unusual posting behaviours of 'hyperactive' Twitter users was found to obscure the voices of the majority of voters (K-C. Yang et al., 2022). This limitation influences how faithfully our sampling method can be considered to capture the views of both groups of individuals we include in our study: autistic self-advocates and parents of autistic children. Our inclusion criteria mandate that users had both Tweeted about autism and included an autism-relevant self-description in their biography, indicating their likely involvement in online autism advocacy. As such, Tweets sampled from these users may include patterns of language that are not as prevalent outside of online advocacy circles.

Future work can address concerns around the representativeness of sampling methods by using techniques adapted from the minimal intervention literature (Pan, 2006; van der Beek et al., 2021). Minimal participation refers to the process through which participants are contacted for necessary information while safeguarding against high non-response rates by requiring minimal time and resource investment. As it pertains to this study, minimal participation may involve sending Twitter users, whom the predictive model classifies as autistic individuals, a message using Twitter API's automated messaging function, requesting their voluntary, compensated and secure participation in an abridged version of a suitable autism inventory, such as the Autism Behavior Inventory (Bangerter et al., 2017), or the Adult Executive Functioning Inventory (Holst & Thorell, 2018). Whether or not the sample can be deemed adequately representative can be determined by comparing the results of the survey to data from large samples of autistic adults (Henninger & Taylor, 2013; Howlin & Moss, 2012; Magiati et al., 2014). Similarly, Twitter users identified as autistic self-advocates or parents of autistic children may be asked to report their views, beliefs and behaviours surrounding autism advocacy, which can be compared to previous surveys conducted outside of the online advocacy context to quantify and account for the non-representativeness introduced by the sampling method.

Limitations of imbalanced classification

The second limitation relates to the use of an automatic classification procedure to generate corpora for further language model training. All automatic predictions introduce the risk of incorrect classification. Especially prone to classification error are predictive models trained on

imbalanced datasets, in which one class far outnumbered the others, yielding predictions that are biased towards the majority class (Sun et al., 2009). In this study, Twitter users who do not openly identify as autistic or parents of autistic children are over-represented in the training data, resulting in poorer predictive performance for the two populations of interest (more details in Supplementary Materials A.1.2). Such a pattern of classification error may pose a threat to the internal validity of the study. Incorrectly sorting a user that belongs to the minority class into the majority sample, for instance, may taint the common baseline from which all word embeddings across the four corpora are derived, leading to potentially inaccurate or misleading estimates of semantic similarity between organisations and individuals. Fortunately, many courses of action can be taken to boost classification accuracy (Burnaev et al., 2015; Maldonado et al., 2022; Shamsudin et al., 2020), including techniques to produce higher-fidelity predictions by improving the quality of the training data itself.

The present methodology includes several measures to address classification error, such as undersampling the majority class of non-autistic, non-parent users and generating new, synthetic examples of the minority class (Chawla et al., 2002), bolstering the model's ability to distinguish between the classes (for more details on undersampling and oversampling, see Supplementary Materials A.1.2). The most effective data augmentation technique, however, is simply to retrieve more real-life examples of minority classes. By supplying the predictive model with new training data in which minority classes are more evenly represented, instances of incorrect classification are reduced. An additional benefit of this technique is that the organic variability of user biographies is preserved, which acts as a deterrent against over-fitting in the training stage, allowing the model to produce accurate predictions for biographies external to the training set.

Conclusion

Set against the backdrop of tensions in autism advocacy, this study explored the alignment of language used by autism-centred organisations and the individuals over whom they may have representational duties. By comparing language use in the Twitter expressions of organisations and individuals, the study provided evidence to support criticisms of partial representation that have been faced by both autism self-advocacy groups and autism charities. To our knowledge, this study is the first computational investigation of a conflict that has largely been characterised by qualitative and opinion-based insights. In future applications, the study's novel methodology and findings may serve as a useful framework to guide organisations and individuals seeking to redress the long-standing conflict in autism advocacy.

Acknowledgements

We would like to thank Ren Palmer for the substantial help with annotating user biographies, and to all graduate students at the Geoff Bird Lab at the University of Oxford for their help with sourcing autism-relevant keywords. We are also grateful to Brennan Delattre, Graham Reid, Jirko Rubek and Damien van de Berg for their helpful comments on the study design and article.

Author Contributions

C.A. was involved in the conceptualisation, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualisation, writing – original draft, writing – review & editing. G.B. was involved in the conceptualisation, investigation, methodology, supervision, writing – original draft.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethics review

Prior to undertaking the research detailed in the article, the authors submitted an ethics review request to an ethics review board at their home institution. The research was determined to be exempt from further ethics review as it involved the analysis of publicly available data with no direct interaction with individuals. The authors adhered to all privacy protections and terms of service stipulated by the Twitter API.

ORCID iDs

Canfer Akbulut  <https://orcid.org/0009-0007-5163-4678>
Geoffrey Bird  <https://orcid.org/0000-0002-2310-0202>

Supplemental material

Supplemental material for this article is available online.

Notes

1. In online discourse, organisations not led by autistic people are often broadly referred to as ‘autism charities’, as many of the most widely recognised non-autistic-led groups are registered as such. However, we wish to avoid this designation, as it carries legal and tax status implications. Instead, we distinguish between the two kinds of organisations considered in this study based only on one criterion: whether an organisation is self-designated as a self-advocacy group or not. This terminology is chosen because it reflects the most accessible and directly operationalisable distinction between these two groups, and is not meant to imply that self-advocacy groups are the only legitimate or defensible unit of organisation in autism advocacy.
2. <https://osf.io/vkqx2>
3. <https://github.com/canferakbulut/TWITAUT>

References

- [@autisticadvocacy]. (2014). ‘*actuallyautistic*’ is a tag for people who are themselves autistic to post in and draw support [post]. Tumblr.
- [@AutSciPerson]. (2021). #saynotooba [post]. Twitter.
- Autistic Self Advocacy Network. (2009). *Horrorific autism speaks ‘I am autism’ ad transcript*.
- Autistic Self Advocacy Network. (2016). *Before you donate to autism speaks, consider the facts*.
- Autistic Self Advocacy Network. (2022). #stoptheshock: *The judge rotenberg center, torture, and how we can stop it*.
- Bangerter, A., Ness, S., Aman, M. G., Esbensen, A. J., Goodwin, M. S., Dawson, G., Hendren, R., Leventhal, B., Khan, A., Opler, M., Harris, A., & Pandina, G. (2017). Autism behavior inventory: A novel tool for assessing core and associated symptoms of autism spectrum disorder. *Journal of Child and Adolescent Psychopharmacology*, 27(9), 814–822.
- Benjamin, E., Ziss, B. E., & George, B. (2020). Representation is never perfect, but are parents even representatives? *The American Journal of Bioethics*, 20(4), 51–53.
- Bertilsdotter Rosqvist, H., Brownlow, C., & O’Dell, L. (2015). ‘An association for all’—notions of the meaning of autistic self-advocacy politics within a parent-dominated autistic movement. *Journal of Community & Applied Social Psychology*, 25(3), 219–231.
- Bowker, N. (2008). Participating in the world (wide web): Social connections for people with disabilities. *Remote Relationships in a Small World*, 41, 98.
- Bowker, N., & Tuffin, K. (2004). Using the online medium for discursive research about people with disabilities. *Social Science Computer Review*, 22(2), 228–241.
- Burnaev, E., Erofeev, P., & Papanov, A. (2015). Influence of resampling on accuracy of imbalanced classification. In A. Verikas, P. Radeva, & D. Nikolaev (Eds.), *Eighth international conference on machine vision (ICMV 2015)* (Vol. 9875, pp. 423–427). SPIE.
- Buturovic, Z. (2020). Parents of autistic children are deserving of support. *American Journal of Bioethics*, 20(4), 54–55.
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186.
- Caselles-Dupré, H., Lesaint, F., & Royo-Letelier, J. (2018). *Word2vec applied to recommendation: Hyperparameters matter*. In Proceedings of the 12Th ACM Conference on Recommender Systems, Vancouver, BC, Canada, (pp. 352–356). Washington, USA: SPIE. <https://doi.org/10.1145/3240323.3240377>
- Chapman, R., & Veit, W. (2020). Representing the autism spectrum. *The American Journal of Bioethics*, 20(4), 46–48.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1), 39–67.
- Des Roches Rosa, S. (2013). *Why did amy s.f. lutz attack the neurodiversity movement? Thinking Person’s Guide to Autism*.
- Dockrell, J. E., Ricketts, J., Charman, T., & Lindsay, G. (2014). Exploring writing products in students with language impairments and autism spectrum disorders. *Learning and Instruction*, 32, 81–90.

- Dockrell, J. E., Ricketts, J., Palikara, O., Charman, T., & Lindsay, G. (2012). *Profiles of need and provision for children with language impairments and autism spectrum disorders in mainstream schools: A prospective study*. Department for Education.
- Elster, N., & Parsi, K. (2020). Like autism, representation falls on a spectrum. *The American Journal of Bioethics*, 20(4), 4–5.
- Fellowes, S. (2020). Additional challenges to fair representation in autistic advocacy. *The New Bioethics*, 20(4), 44–45.
- Fleischmann, A. (2005). The hero's story and autism: Grounded theory study of websites for parents of children with autism. *Autism*, 9(3), 299–316.
- Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 115(16), E3635–E3644.
- Glumbić, N., Đorđević, M., & Brojčin, B. (2022). *Self-Advocacy* (pp. 215–229). Springer International Publishing.
- Goldberg, Y., & Levy, O. (2014). *word2vec explained: Deriving Mikolov et al.'s negative-sampling word-embedding method*. arXiv Preprint.
- Graff Zivin, J., & Small, A. (2005). A modigliani-miller theory of altruistic corporate social responsibility. *The BE Journal of Economic Analysis & Policy*, 5(1), 1–21.
- Hall, C. M., Culler, E. D., & Frank-Webb, A. (2016). Online dissemination of resources and services for parents of children with autism spectrum disorders (asds): A systematic review of evidence. *Review Journal of Autism and Developmental Disorders*, 3(4), 273–285.
- Haney, J. L. (2018). Reconceptualizing autism: An alternative paradigm for social work practice. *Journal of Progressive Human Services*, 29(1), 61–80.
- Henninger, N. A., & Taylor, J. L. (2013). Outcomes in adults with autism spectrum disorders: A historical perspective. *Autism*, 17(1), 103–116.
- Holst, Y., & Thorell, L. B. (2018). Adult executive functioning inventory (ADEXI): Validity, reliability, and relations to ADHD. *International Journal of Methods in Psychiatric Research*, 27(1), Article e1567.
- Howlin, P., & Moss, P. (2012). Adults with autism spectrum disorders. *The Canadian Journal of Psychiatry*, 57(5), 275–283.
- Hur, W.-M., Moon, T.-W., & Kim, H. (2020). When and how does customer engagement in CSR initiatives lead to greater CSR participation? the role of CSR credibility and customer–Company identification. *Corporate Social Responsibility and Environmental Management*, 27(4), 1878–1891.
- Jaeger, P. T. (2022). *Disability and the internet* [In Disability and the Internet]. Lynne Rienner Publishers.
- Jamali, D., & Mirshak, R. (2007). Corporate social responsibility (CSR): Theory and practice in a developing country context. *Journal of Business Ethics*, 72, 243–262.
- Joseph, M. (2011). *New wandering code approved*. Autism Jabberwocky.
- Kapp, S. K., Gillespie-Lynch, K., Sherman, L. E., & Hutman, T. (2013). Deficit, difference, or both? autism and neurodiversity. *Developmental Psychology*, 49(1), 59.
- Koegel, L. K., Bryan, K. M., Su, P. L., Vaidya, M., & Camarata, S. (2020). Definitions of nonverbal and minimally verbal in research for autism: A systematic review of the literature. *Journal of Autism and Developmental Disorders*, 50(8), 2957–2972.
- Kofmel, E. (2019). *ICT accessibility for autistic persons and self-advocacy online and off*. Text of the Invited Oral Presentation by Erich Kofmel, President of Autistic Minority International at the JCA-AHF Meeting at ITU Headquarters in Geneva.
- Krcek, T. E. (2013). Deconstructing disability and neurodiversity: Controversial issues for autism and implications for social work. *Journal of Progressive Human Services*, 24(1), 4–22.
- Kusner, M., Sun, Y., Kolkin, N., & Weinberger, K. (2015). *From word embeddings to document distances*. In International Conference on Machine Learning, Lille, France (pp. 957–966).
- Lawson, M. A., Martin, A. E., Huda, I., & Matz, S. C. (2022). Hiring women into senior leadership positions is associated with a reduction in gender stereotypes in organizational language. *Proceedings of the National Academy of Sciences*, 119(9), Article e2026443119.
- Lin-Hi, N., Hörisch, J., & Blumberg, I. (2015). Does CSR matter for nonprofit organizations? Testing the link between CSR performance and trustworthiness in the nonprofit versus for-profit domain. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 26(5), 1944–1974.
- Lock, I., & Seele, P. (2017). Measuring credibility perceptions in CSR communication: A scale development to test readers' perceived credibility of CSR reports. *Management Communication Quarterly*, 31(4), 584–613.
- Lutz, A. (2013). *Is the neurodiversity movement misrepresenting autism?* Slate.
- Lutz, A. (2021). *People with disabilities want a wide range of options*. Bloomberg.
- Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., Park, B. Y., Snyder, N. W., Schendel, D., Volk, H., Windham, G. C., & Newschaffer, C. (2017). The changing epidemiology of autism spectrum disorders. *Annual Review of Public Health*, 38, 81–102.
- Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: A systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review*, 34(1), 73–86.
- Maldonado, S., Vairetti, C., Fernandez, A., & Herrera, F. (2022). Fw-smote: A feature-weighted oversampling approach for imbalanced classification. *Pattern Recognition*, 124, 108511.
- McCoy, M. S., Liu, E. Y., Lutz, A. S. F., & Sisti, D. (2020). Ethical advocacy across the autism spectrum: Beyond partial representation. *The American Journal of Bioethics*, 20(4), 13–24.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 2, 9.
- Ne'eman, A. (2011). *Action alert: Tell the CDC 'no' on abuse-enabling 'wandering' code!* Autistic Self-Advocacy Network.
- Ne'eman, A. (2021). *People with disabilities need services in the community*. Bloomberg.

- Ne'eman, A., & Bascom, J. (2020). Autistic self advocacy in the developmental disability movement. *The American Journal of Bioethics*, 20(4), 25–27.
- National Institute of Neurological Disorders and Stroke. (2018). *Autism spectrum disorder fact sheet*.
- Ortega, F. (2009). The cerebral subject and the challenge of neurodiversity. *Biosocieties*, 4(4), 425–445.
- Pan, W. (2006). Proactive telephone counseling as an adjunct to minimal intervention for smoking cessation: A meta-analysis. *Health Education Research*, 21(3), 416–427.
- Parsloe, S. M. (2015). Discourses of disability, narratives of community: Reclaiming an autistic identity online. *Journal of Applied Communication Research*, 43(3), 336–356.
- Pellicano, E., & den Houting, J. (2022). Annual research review: Shifting from 'normal science' to neurodiversity in autism science. *Journal of Child Psychology and Psychiatry*, 63(4), 381–396.
- Rehurek, R., & Sojka, P. (2011). Gensim–Python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, 3(2).
- Reinke, J. S., & Solheim, C. A. (2015). Online social support experiences of mothers of children with autism spectrum disorder. *Journal of Child and Family Studies*, 24(8), 2364–2373.
- Richardson, J. P., & Sharp, R. R. (2020). Meaningful fissures: The value of divergent agendas in patient advocacy. *The American Journal of Bioethics*, 20(4), 1–3.
- Richman, K. A. (2020). Neurodiversity and autism advocacy: Who fits under the autism tent? *American Journal of Bioethics*, 20(4), 33–34.
- Rose, V., Trembath, D., Keen, D., & Paynter, J. (2016). The proportion of minimally verbal children with autism spectrum disorder in a community-based early intervention programme. *Journal of Intellectual Disability Research*, 60(5), 464–477.
- Rosenblatt, A. (2018). Autism, advocacy organizations, and past injustice. *Disability Studies Quarterly*, 38(4).
- Sanchack, K. E., & Thomas, C. A. (2016). Autism spectrum disorder: Primary care principles. *American Family Physician*, 94(12), 972–979.
- Shamsudin, H., Yusof, U. K., Jayalakshmi, A., & Khalid, M. N. A. (2020). *Combining oversampling and undersampling techniques for imbalanced classification: A comparative study using credit card fraudulent transaction dataset*. In 2020 IEEE 16Th International Conference on Control & Automation (ICCA), Singapore (pp. 803–808).
- Siegel, B. (2020). Hope springs eternal: Pitfalls of partial representation by advocates who only want the best. *The American Journal of Bioethics*, 20(4), 38–40.
- Sun, Y., Wong, A. K., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(4), 687–719.
- Tabor, H. K. (2020). Civic engagement, autism and deliberative democracy: Prioritizing the inclusion of marginalized perspectives. *The American Journal of Bioethics*, 20(4), 41–43.
- Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research*, 6, 468–478.
- Trevisan, F. (2016). *Disability rights advocacy online: Voice, empowerment and global connectivity* (pp.81–102). Routledge.
- van der Beek, S., Bellhäuser, H., & Hertel, S. (2021). Do minimal interventions increase participation rates in voluntary online training at high school? *Psychology Learning & Teaching*, 20(3), 348–363.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R. . . . SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17, 261–272.
- Wilson, J. (2013). *No more – A letter to Suzanne Wright*. A Diary of a Mom.
- Yang, J., Basile, K., & Letourneau, O. (2020). The impact of social media platform selection on effectively communicating about corporate social responsibility. *Journal of Marketing Communications*, 26(1), 65–87.
- Yang, K.-C., Hui, P.-M., & Menczer, F. (2022). How twitter data sampling biases us voter behavior characterizations. *PeerJ Computer Science*, 8, Article e1025.
- Zuber, W. J., & Webber, C. (2019). Self-advocacy and self-determination of autistic students: A review of the literature. *Advances in Autism*, 5(2), 107–116.