Review of cluster analysis of phenotypic data in Autism Spectrum Disorders: distinct subtypes or a severity gradient model?

Christine K. SyriopoulouDelli, Department of Educational and Social Policy, University of Macedonia

csyriop@gmail.com

Elpis Papaefstathiou, Department of Educational and Social Policy, University of Macedonia

papaelpi@yahoo.gr

Running Head

Review of classifying Autism Spectrum Disorders

Corresponding author:

Christine K. SyriopoulouDelli, Postdoc, PhD, MEd

Assistant Professor: Autism Spectrum Disorders

Department of Educational and Social Policy

University of Macedonia, Thessaloniki

156 EgnatiaSt.

54336 Thessaloniki

Mobile phone +030 6938704339

csyriop@gmail.com

Abstract

Background: Individuals with autism spectrum disorder (ASD) form a heterogeneous group, posing a challenge for clinical definition. Additional problems regarding the diverse clinical presentation arise from changes in diagnostic criteria according to the latest Diagnostic and Statistical Manual of Mental Disorders (DSM-5), with exclusion of individuals who met earlier criteria or inclusion of more than previously.

Objectives: To investigate studies that have attempted to reduce the heterogeneity of ASD based on cluster analysis of phenotypic data and to clarify whether ASD should be interpreted as "a unitary spectrum", with a severity gradient, or defined by distinct subtypes. This will allow better understanding of the disorder with implications for its treatment and prognosis.

Methods: A literature search was made through PubMed, Researchgate and Google Scholar for studies of ASD populations. In addition, reference lists from identified studies were reviewed.

Results: Only 10 studies were found that dealt with the heterogeneity of ASD and its different subtypes, based on the review prerequisites. Most of the studies appear to support the existence of subtypes within ASD, but it remains unclear whether these are considered as different specific subtypes with characteristic profiles of symptoms or as a part of a severity gradient across symptom domains.

Conclusions: Drawing definitive conclusions from the published studies about the nature of ASD is difficult, due to the fundamental methodological differences among the studies and their inconsistent findings. This review shed light on a number of discrepancies regarding the current classification of ASD. However, future research will be necessary to provide a more definite answer on the question of a definition based on separate diagnostic subtypes or on a severity gradient by including larger

samples that are followed longitudinal and by applying better diagnostic system and choosing the appropriate variables.

Key words: Autism spectrum disorder, Cluster analysis, DSM-5, Subtypes, Severity Gradient

Introduction

Individuals with autism spectrum disorder (ASD) form a heterogeneous group with variable clinical presentation, including a wide range in age of onset, cognitive level and the severity of specific behaviours such as social skills, communication and language ability (Persico & Bourgeron, 2006; Cholemkery et al., 2016). This heterogeneity has received considerable research attention and poses a challenge for the definition of clinical populations (Brennan et al., 2015).

Additional problems regarding definition are related to changes in the diagnostic criteria. In previous editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM), and specifically DSM-III, DSM-III-R, DSM-IV and DSM-IV-TR, (Waterhouse, 1992; American Psychiatric Association, 2000),the category pervasive developmental disorders (PDD) included various different diagnostic subgroups, such as autism, Asperger's syndrome and pervasive developmental disorder-not otherwise specified (PDD-NOS). According to the most recent edition, DSM-5(American Psychiatric Association, 2013), the currently used term ASD is characterized by two core dimensions, namely social reciprocity and restricted or stereotyped behaviours or interests, which vary in quality, quantity and severity in each individual. Because poor reliability was reported in how subtypes were assigned to each individual, DSM-5 removed the earlier different diagnostic subcategories (Grzadzinski et al., 2013).

Diagnosis according to subtypes may have appeared to be clearer in the categorical system, as in the previous DSM editions, the emphasis was given on the different diagnostic subgroups separately, but it impeded emphasis on the heterogeneity of ASD because in the DSM-5 there is only one spectrum where the wide range of phenotypic variation of individuals makes a very heterogeneous clinical entity. For this reason, the term ASD is used as an umbrella to cover all the previous diagnostic categories (American Psychiatric Association, 2000).

Following this reorganization in the diagnostic criteria of ASD, however, certain concerns have emerged. Firstly, the new system may exclude from a diagnosis some individuals who previously met the criteria for a diagnostic subtype. Secondly, it appears that the concept of PDD has been broadened to the point where, in some cases, individuals with very different types of difficulties are included in the same diagnostic category. Specifically, in DSM-5, the term "spectrum" covers the heterogeneity and wide range of phenotypic variation of individuals with ASD. Symptoms are described as falling along a continuum, with the more severe symptoms at one end of the spectrum and the less severe at the other end. The continuum as it is defined by the DSM-5 diagnostic criteria is related to the severity of both social communication impairments and restrictive repetitive patterns of behaviour. In practice, this interpretation is difficult when applied to individuals who vary in terms of quality and quantity in these two domains (Buitelaar et al., 1999; Walker et al. 2004).

As all individuals are classified within a single diagnostic group we are urged to move to the notion of a spectrum disorder and to adopt a dimensional model rather than the previous categorical system (Allison et al., 2012), although it has not been clarified whether ASD can be interpreted by "a unitary spectrum" model or still needs

to be defined by distinct subtypes. Specifically, the unitary spectrum model based on the DSM-5 refers to a continuum that includes the domains of both social communication and restrictive repetitive patterns of behaviour. So, according to these the different subgroups present differences in quantity in these two domains, whereas concerning the distinct subgroups theory there are qualitative differences among them and each one of the subgroups present a different pattern of symptoms.

Many studies have focused on stratifying groups in order to improve phenotypic homogeneity, and some of the first attempts to reduce heterogeneity were made long before this dilemma was posed (Lotter, 1974).The main purpose of this approach was to determine the etiology and physiology of the various different ASD subtypes, and to individualize the appropriate forms of intervention.

The first approach was the application of taxonomic systems based on clinical impressions (Matson &Mulick, 1991), but later studies attempted to reduce heterogeneity using statistical methodology like factor analysis and cluster analysis. According to Snow, Lecavalier and Houts(2008), the use of factor analysis can only result in associations between some variables and underlying latent factors. It is an explorative analysis which groups similar variables into dimensions, simplifying the data by reducing the dimensions of the observations and helping to identify underlying factors. In most studies, homogeneity has typically been defined by a simple variable, but ASD groups usually differ in many dimensions; variation in levels of intellectual functioning, language and communication all contribute to the heterogeneity of autism. Many attempts at classification were based on subtyping by ASD behaviour and symptoms either separately or in combination (Wing & Gould, 1979).

The studies of Rescorla (1988) and Siegel and colleagues (1986) were based on empirical findings using the statistical method of cluster analysis. These studies were considered innovative as for the first time cluster analysis was applied in order to investigate the nature of ASD. Although there were methodological flaws in these studies, their contribution was significant. Boucher, Lewis and Collis (1998) suggested that it would be more appropriate to describe children based on a multidimensional rather than a categorical view of autism. Cluster analysis can be applied to investigate empirically derived diagnostic differences between the subtypes and the phenotypic characteristics of groups, reducing the heterogeneity of ASD. Cluster analysis is used to group data according to the characteristics they possess, and the grouping is based on the similarity of factors, with no differentiation between dependent and independent variables. The main objective of cluster analysis is to address the heterogeneity in each dataset, grouping items that share common characteristics.

Although cluster analysis and factor analysis are both used for simplifying data, and may appear similar, they also differ in many ways. Firstly, the objectives are different; the objective of cluster analysis is to divide the observations into homogeneous and distinct groups, while factor analysis explains the homogeneity of the variables resulting from the similarity of values. Factor analysis is suitable for simplifying complex models by reducing large sets of variables to a much smaller set of factors, whereas cluster analysis is suitable for classifying objects based on certain criteria; cluster analysis attempts to group cases whereas factor analysis attempts to group features. So, cluster analysis groups individuals with ASD according to their similarity for a wide range of variables. The hypothesis behind the preference for this method is based on the fact that cluster analysis results either in a single cluster,

meaning that there are no subgroups in ASD, or in more than one cluster differing in the quantity of domains, but not in the quality of different traits (Palmer et al., 2015).

In an attempt to contribute to this debate on whether ASD consists of different subgroups or is defined by a severity gradient model, the current review focused on the results of the various efforts that researchers have made over the last 25 years to classify ASD. To our knowledge, this is the only review of the literature ranging from the first attempts at reducing heterogeneity, even before the new diagnostic criteria were introduced, up to the latest studies based on the DSM-5 criteria. Our objective was to provide a better understanding of the different profiles of ASD and the current classification.

Method

Identification of studies included both electronic and manual search and selection strategies. A literature search was conducted n PubMed, ResearchGate and Google Scholar to locate studies published between 1994 and 2018 that examined the heterogeneity of ASD and its various different subtypes by entering the terms: autism, ASD, subtype, subgroups, cluster analysis, classification. The reference lists from the studies that were identified were also reviewed. In the first screening, titles and abstracts were evaluated according to the sample characteristics of the study, criteria of classification and statistical approach. This yielded 19 studies of which the full text was reviewed. After application of the inclusion and exclusion criteria for this review, only 10 articles were found suitable for inclusion. Specifically, the inclusion prerequisites were:

1. Cluster analysis was applied for statistical analysis of the data.

2. The cluster analysis was based on data from multiple sources and domains concerning the characteristics of individuals with ASD.
3. The study sample consisted of individuals with a variety of profiles within the spectrum of the disorder.

Studies that applied a statistical method other than cluster analysis to address heterogeneity (e.g., factor analysis, latent class analysis) were excluded. In addition, studies in which clustering was based exclusively on a single source of data, for example from magnetic resonance imaging (MRI) or on only a specific domain of assessment (e.g., sensory processing), were not included in this review. Studies that used a combination of behaviours, traits and domains were preferred, in order to capture the broad phenotypic heterogeneity of ASD and differing profiles of the individuals with ASD. For the same reason studies in which the samples included only individuals in a single diagnostic category based on an earlier DSM edition (such as high-functioning autism, Asperger's syndrome or PDD-NOS) were not included in this review.

Results

Table 1 lists the 8 studies that met the inclusion criteria, showing the authors, publication year, sample size and characteristics, the diagnostic criteria used, the variables used for the cluster analysis, the number of clusters derived, and which support a severity gradient model and which the use of different diagnostic subtypes.

INSERT TABLE 1 SOMEWHERE AROUND HERE

The First Empirically Based Attempts

Eaves, Ho and Eaves (1994) conducted one of the first studies that attempted to highlight the important differences among individuals with autism. As, according to the DSM-III-R criteria in effect at the time of their study, there was a "fogginess" in the diagnosis given to a large proportion of children with what is now referred to as ASD, the research team decided to study a large group with a range of autistic features. This was considered a pioneering study because of its large sample and multidisciplinary data, as the previous reports had been based only on one kind of data. Specifically, the objective of this study was to identify possible subgroups based on behavioural and cognitive data and medical and family histories. Cluster analysis identified four separate subtypes that differed from each other in some behavioural and cognitive aspects, and in their social skills, communication and sensory preoccupations. The first subgroup consisted of typically autistic children with mild mental disability and the second subgroup of autistic with moderate to severe mental disability. In the third and fourth subgroups there were higher functioning children; specifically in the third subgroup the children presented the least autistic traits and those in the fourth subgroup were the most highly functioning subgroup, but with many differences in all the domains that were assessed. The four subtypes differed in some behavioural characteristics such as subgroup one were aloof, with abnormal nonverbal communication and language, motor stereotypies or sensory preoccupations. Subgroup two were aloof and nonverbal and indulged in motor stereotypies and sensory preoccupations. Subgroup three had mainly sensory preoccupations and restricted interests whereas subgroup four had abnormal language with motor stereotypies and restricted interests. These are some of the characteristics that make these subgroups to differ in their symptoms profile.

Although this study was carried out long before the new diagnostic criteria were introduced, the identification of four distinct subgroups in a large sample supports the notion of different clinical subtypes within the autistic spectrum. It was considered of great importance that the DSM-III-R in use at that time could not provide categories for all four subtypes identified in the analysis, meaning that a proportion of children were possibly being left undiagnosed and without appropriate intervention.

Sevin and colleagues (1995) conducted a similar study, the results of which, in general, supported the existence of different homogeneous subgroups. It was the first study in which the cluster analysis was based on standardized, reliable and valid instruments, specifically for individuals within the autistic spectrum. In this study, also, the analysis resulted in four clusters. Specifically, cluster 1 included the highest functioning individuals in terms of social skills and sensory abnormalities, cluster 4 the individuals with severe autism in terms of social and language impairments and of stereotypic/ritualistic behaviour. Clusters 2 and 3 included those with intermediate levels of difficulties. Regarding these results, cluster 2 could be identified as "mild autism" with mild levels of stereotypies and sensory abnormalities and cluster 3 as "moderate autism" with moderate social and language impairments but by severe levels of sensory abnormalities and stereotypies. Despite the fact that these two clusters appeared similar, distinguishable differences were detected in domains such as sensory processing and self-stimulatory behaviours. The clusters, in this study as well differed mainly in the traits that form the different symptom profiles. Although the study sample was very small, it included the full spectrum of developmental disorders. The results indicated the existence of four different subtypes in terms of

behaviour related to autistic traits, just as the earlier study of Eaves and colleagues (1994).

However later on, in Bitsika's and al. study in 2008 demonstrated the limited contribution concerning the assessment procedures in a categorically system due to their findings children with Asperger and autism might be reclassified into subgroups based on the presence of specific behaviours. This implies that the notion of a 'severity model' would fit better as the distinction between the different subgroups is not so clear. Specifically, they ended up on three different clusters, where the first cluster was the highest functioning subgroup, the second one the moderately functioning and the last one the low-functioning group in terms of skill impairments in communication, social interaction and adaptive behaviour. Their differences referred to the degree in which all the individuals vary in some domains and the severity of social and cognitive impairments. As distinctive patterns of symptoms were not revealed, the study supports the concept of a continuum.

Studies including Longitudinal Examination

The main purpose of the study of Fein and colleagues (1999), was to indentify subtypes in a preschool sample of children with PDD, and the behaviour patterns that best discriminate these subtypes. The study also followed the trajectories of some of the children during school age to verify the longitudinal stability of autistic subtypes. Various statistical methods were used for the detection of subtypes, and cluster analysis was then applied, using social, language, cognitive and behavioural data, which validated the two-subtype differentiation that emerged.

The first group was comprised of the higher functioning children, and the second the lower functioning children with PDD. Although the traits in which the

children in the two groups differed could be considered to fall along a continuum, with scores ranging from low to high, this did not apply to the relationships among the continuous variables, which led to the conclusion that the groups were discontinuous and constitute separate diagnostic categories (i.e., subtypes). For example while verbal IQ, nonverbal IQ and severity of symptoms proved to be continua the relationships among these variables are different between the two groups. Cognitive level proved to be the factor that best predicts the membership in one of the two groups; the children in both groups shared many difficulties, but differed in their cognitive level and the degree of difficulties, with children in the lower-functioning group showing more severe manifestations of symptoms. Although, it was found a longitudinal stability concerning the membership in the two groups, the higher functioning children had greater probability of not appearing PDD by school age and improving in the domains of social interaction, communication and patterns of restrictive behaviour while the lower functioning children remained relatively stable.

Similar findings were reported by Stevens and colleagues (2000), whose cluster analysis study using variables related to expressive and receptive language, nonverbal intelligence and social behaviour, validated the two-group differentiation. Their study included longitudinal examination aimed at identifying the preschool variables that best predict school-age functioning. They showed that in the lower-functioning group, language and social scores either remained stable or drop from preschool to school age, while in the higher-functioning group the scores in all the variables improved. Nonverbal IQ appeared to be the best predictor of subgroup membership at school age. An important finding was that about one third of the children in the high-functioning group at preschool, improved sufficiently to achieve normal scores at school age.

On the contrary, in Wiggins's et al. study (2012) a dimensional view of ASD is supported. The study focuses in toddler population. Three subgroups were emerged in terms of social, communication and intellectual abilities, repetitive behaviours and abnormal sensory response. The first subgroup the 'ASD, mild impairment' had the best communication abilities and adaptive behaviour with few deficits in communication in stereotyped interests and behaviours and low-average intellectual abilities. The second one labelled as 'moderate impairment' had many social and communication deficits, few stereotyped interests and behaviours and mild intellectual disability. The third one 'the severe impairment' showed many social and communication impairments, as well as in repetitive behaviours and abnormal sensory response than the other two subgroups and mild-moderate intellectual disability. These findings support the hypothesis of a continuum of behaviours due to the differences in quantity of deficits among the three subgroups. Concerning the longidutinal findings of this study no significant differences were found after a period of two years except for the fact that 23 children did no longer meet the diagnostic criteria for an ASD.

Attempts at identifying genetically phenotypic subgroups

A popular hypothesis is that the wide variability in clinical manifestations of ASD can be explained by underlying genetic heterogeneity, and thus the identification of more phenotypically homogeneous subgroups may help to elucidate the underlying genetic mechanisms. Spiker and colleagues (2002) investigated the behavioural phenotypic variability of siblings with ASD aiming to find "susceptibility" genes for autism. Many other genetic researchers, over the last two decades, have claimed that there are subgroups of families of children with ASD with distinct underlying etiology

(Folstein et al., 1998; Szatmari et al., 1998) and some have attempted to define these subgroups based on behavioural patterns (Castelloe& Dawson, 1993).Clustering of symptoms is thus a common approach to the definition of subgroups of families.

According to Spiker and colleagues (2002) results*)*, on cluster analysis three clusters emerged. Cluster 1 was labelled mild, cluster 2 moderate and cluster 3 severe ASD. As the symptoms are based on a continuum from mild to severe, it was proposed that autism should be characterized on a severity gradient. Specifically, no evidence of distinct subgroups of ASD based on symptoms was indentified. The higher the nonverbal IQ scores of the children, the less severe impairments they showed. This relationship remained the same even when the influence of verbal ability was removed from the cluster analysis. The only exceptions to the severity gradient model concerned repetitive and ritualistic symptoms, circumscribed interests and compulsions/rituals.

Hu and Steinberg (2009) attempted to indentify subgroups of ASD in a large sample, using multiple clustering methods, and proceeded to gene expression analysis. Cluster analysis resulted in four distinct phenotypic subgroups. One group included individuals with severe language impairment, the second those with milder symptoms across all domains, the third those with higher levels of savant skills, and the fourth those with intermediate severity across all domains. The average of item scores for each group across all domains suggested phenotypic distinction between the groups, although some quantitative differences in severity between the savant and the mild groups were observed across the functional and behavioural domains. This study thus offered support for both the discrete phenotype (subtype) model and the gradient model, by identifying sets of genes that are expressed differentially, either quantitatively or qualitatively among ASD subgroups.

Veatch and colleagues (2014) also attempted to classify variables in order to identify clinical subgroups. They used multiple sources of behavioural data and also biomarker data, such as head circumference (HC). Apart from for the novel use of this kind of data, a further original element of this study was the use of two separete large ASD datasets, in order to replicate and confirm the findings. Their results confirmed a severity-based dichotomy, with the identification of two distinct clusters. Specifically, the larger cluster had more individuals with higher severity scores, and was labelled the "more severe" and the smaller cluster was labelled the "less severe". In a further step, cluster analysis was applied separately to each of the two clusters, resulting in six "subclusters" for the "more severe" and four "subclusters" for the "less severe" cluster.

Classifying ASD based on DSM-5

The important element of the study of Cholemkery and colleagues (2016) was that their cluster analysis was based on the current DSM-5 criteria. The research team used the subscales of the Autism Diagnostic Interview – Revised (ADI-R) algorithm domains which were reorganized to current criteria. This process produced three clusters; cluster 1 included the most severely affected, who showed more deficits in three domains (reciprocal social interaction, communication, stereotyped behaviours, and abnormal development), while cluster 2 included those who showed moderate symptom severity in three of the four domains and minor severity in the domain of stereotyped behaviour. The characteristics of social interaction, communication and stereotyped repetitive behaviour were shown to fall along a continuum, as used in DSM-5. The individuals in cluster 3 showed the least severe symptoms in the three of the four domains, but moderate severity in stereotyped behaviour. Thus, while the

most severely affected individuals were allocated to one cluster, those less severely affected were divided between two clusters with differing symptom profiles. According to these findings, this most recent study provides support for both the existence of distinct subtypes within ASD and the severity gradient model.

Discussion

Many researchers have attempted to identify homogenous and clinically meaningful subtypes of ASD (Ring et al., 2008; Constantino et al., 2004), in the hope of shedding light on the corresponding underlying factors. The theory was that this would provide an indication of the prognosis and the appropriate types of intervention and treatment, as behaviourally distinct diagnostic subtypes may be associated with specific responses to specific interventions.

Cluster analysis is an empirically derived statistical method that is data driven and groups individuals based on their common characteristics. The studies reviewed here that applied cluster analysis to samples of individuals with ASD identified between one and four clusters, namely subtypes of ASD, which were differentiated mainly by symptom severity, in either quality or quantity. Usually, these subtypes were labelled as severe, moderate and mild ASD (e.g., Eaves et al. 1994; Sevin et al., 1995). In the studies where only two clusters emerged, the data indicated that the clusters differed in symptom severity and usually are labeled as high- or low-functioning (Fein et al., 1999; Stevens et al., 2000). Cohen, Paul, and Volkmar (1987) and Tsai (1992) had suggested earlier that high- and low-functioning autism should be considered as different diagnostic entities.

Ring and colleagues (2008) later proposed a severity gradient model, in which the differences among individuals with ASD are quantitative rather than qualitative,

and Constantino and colleagues (2004) also supported the notion of a continuum in terms of a range of deficits, giving weight to the severity gradient model adopted in DSM-5.

Some studies support both the existence of different subtypes and the severity gradient model (Hu and Steinberg, 2009; Cholemkery et al., 2016). Although most studies to date appear to support the existence of different subtypes within ASD it is still unclear whether the study results indicate different subtypes separate from a continuum, or if the subtypes might be based on the continuum of a severity gradient, labelled with subcategories. Probably, the notions of a 'severity gradient' and 'distinct subtypes' aren't mutually exclusive, but using both might be needed to understand heterogeneity in ASD. Drawing definitive conclusions is difficult, due to the small number of studies and their difference in samples and methodology. In order to answer this question, the empirical subclassification resulting from these studies needs to be replicated and validated. Based on this review, we concluded that there are specific issues that need to be taken into consideration in the design of future studies.

Firstly, the sample size is a crucial factor, as larger samples with a broad range of ages and clinical characteristics provide the chance of including individuals representing a wider range within the spectrum from the less to the more severely affected, enhancing the validity of generalization of the study results. Small sample sizes and limited age groups were regarded as restrictive for the research findings, as was exclusion of certain categories. The studies, for example, of Cholemkery and colleagues (2016) in which non-verbal children were excluded, and Spiker (2002) which did not include individuals from the whole spectrum, resulted in limited power in the generalization of their findings.

A further positive step in future research would be the following of samples from childhood into adolescence and adulthood, which would enable validation of the different trajectories of the subtypes and investigation of the effect of intervention. In the studies that included longitudinal examination (Fein et al., 1999; Stevens et al., 2000), unfortunately no data on intervention were analyzed between evaluations and it could not be determined whether specific factors were responsible in the case of improvement of performance.

All assessments should be carefully implemented using the gold standard diagnostic system, to ensure the internal validity of each study and enable comparison between studies. Some studies included the data from one primary instrument while others used a combination. These methodological differences may have resulted in inconsistent findings. The modern diagnostic methods need to be applied more strictly. ADI-R scores, for example, provide the most widely available phenotypic data on behavioural and functional ASD symptoms for the majority of children with ASD. Identification of possible subgroups within ASD is made by evaluating behavioural data from as many as possible multiple sources (Georgiades et al., 2013).Veatch and colleagues (2014) used a wide range of measurement variables to identify clusters. They applied a novel approach of multivariate statistical analysis including both behavioural information from multiple sources and quantitative data, such as macrocephaly. Cholemkery and colleagues (2016) added cognitive level, child psychiatric variables and comorbidities to the cluster analysis and this evaluation of a wide range of data potentially enabled more accurate phenotypic definitions. It should be kept in mind, however, that some instruments such as ADI-R were developed as categorical diagnostic instruments and that this might affect the clusters that emerge and limit the possibility of identification of a severity gradient model.

An issue of paramount importance is the choice of the variables on which the cluster analysis is based. It is acknowledged that selection of different variables for clustering may result in different cluster solutions, but the inclusion of too few may bias the result. The variables should be commonly accepted factors in the core impairments in autism. Most researchers agree on the importance of IQ when defining subtypes (Garmezy & Rutter, 1983; Witwer & Lecavalier, 2008), and so inclusion of IQ is necessary in this kind of study. In the study of Spiker and colleagues (2002) it is possible that distinct qualitative clusters might have emerged if a wider range of variables had been included in the cluster analysis. The selection of the variables for clustering should be based on firm theoretical rationales and utilize the current most widely accepted definitions of autism. For this reason, future research should be focused on the DSM-5 criteria. In addition, use of the same statistical methodology will help to clarify the definition of ASD. Although all the studies that were reviewed had applied cluster analysis, they did not follow exactly the same methodology. For example, Fein and colleagues (1999) used first regression analysis for classification and then cluster analysis to confirm the groups, and Veatch and colleagues (2014) used 7 different clustering methods. Based on the experience of these earlier studies, it would probably be more productive for studies to follow the same pattern of statistical analysis. Also, we should keep in mind that the clusters that are found could be in part a faction of test error due to the systematic differences between tests. This could result in clusters being specific to test used.

The criteria change in DSM-5 made because of the lack of empirical data supporting the hypothesis of distinct disorders within ASD, but this should also highlight the need for further studies (Frazier et al. 2012; Snow et al. 2009). How DSM-5 can improve sample definition and help us to identify subtypes of ASD is still

under question. The possibility of distinct ASD groups based on specific patterns of sensory abnormalities and social-communication profiles should be explored (Baker et al., 2008), and factors such as cognitive ability, language ability, comorbid psychopathology and age of onset should be included in future studies.

This review had a number of limitations. It was conducted to highlight the most important information from studies that attempted to define ASD, and the inclusion criteria were strict.

Perhaps, the most significant concern was the generally restricted quantity of available research, based on our prerequisites. No two studies followed exactly the same methodology. The use of different samples, with different age ranges, diagnostic methods and criteria and selection of variables made it difficult to make any definite conclusions concerning the traits of individuals with ASD. This review reviewed only studies applying the cluster analysis method. As studies using different methods of statistical analysis were not included, it is possible that important results may have been excluded. This review, in spite of the limitations, shed light on a number of discrepancies regarding the classification of ASD. Further research with other samples will be needed for validation of the study results. If consistent findings are confirmed by replication and follow-up we could possibly make some conclusions about the nature of ASD, and answer the question of subtypes or severity gradient or both. This will allow better understanding of the disorder and its course and symptomatology, with implications for its treatment and prognosis. Children deserve nothing less than the best possible education, but that does not mean that it must be the same for all. We should first be able to define the characteristics of children with ASD, before deciding what is best for each child (Brennan et al., 2014).

References

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(2), 202-212. DOI:10.1016/j.jaac.2011.11.003

American Psychiatric Association, & American Psychiatric Association (2000). DSM-IV-TR: Diagnostic and statistical manual of mental disorders, text revision. *Washington, DC: American Psychiatric Association*, *75*, 78-85.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.doi.org/10.1176/appi.books.9780890425596

Baker, A. E., Lane, A., Angley, M. T., & Young, R. L. (2008). The relationship between sensory processing patterns and behavioural responsiveness in autistic disorder: A pilot study. *Journal of Autism and Developmental Disorders*, *38*(5), 867-875. DOI:10.1007/s10803-007-0459-0

Bitsika, V., Sharpley, C. F., & Orapeleng, S. (2008). An exploratory analysis of the use of cognitive, adaptive and behavioural indices for cluster analysis of ASD subgroups. *Journal of Intellectual Disability Research*, *52*(11), 973-985. DOI: 10.1111/j.1365-2788.2008.01123 .x

Boucher, J., Lewis, V., & Collis, G. (1998). Familiar face and voice matching and recognition in children with autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(2), 171-181.

Brennan, L., Barton, M., Chen, C. M., Green, J., & Fein, D. (2015). Detecting subgroups in children diagnosed with pervasive developmental disorder-not otherwise specified. *Journal of Autism and Developmental Disorders*, *45*(5), 1329-1344.DOI: 10.1007/s10803-014-2295-3

Buitelaar, J. K., Van der Gaag, R., Klin, A., &Volkmar, F. (1999). Exploring the boundaries of pervasive developmental disorder not otherwise specified: Analyses of data from the DSM-IV autistic disorder field trial. *Journal of Autism and Developmental Disorders*, *29*(1), 33-43. doi.org/10.1023/A:102596653.

Castelloe, P., & Dawson, G. (1993). Subclassification of children with autism and pervasive developmental disorder: a questionnaire based on Wing's subgrouping scheme. *Journal of Autism and Developmental Disorders*, *23*(2), 229-241. doi.org/10.1007/BF01046217.

Cholemkery, H., Medda, J., Lempp, T., & Freitag, C. M. (2016). Classifying autism spectrum disorders by ADI-R: subtypes or severity gradient?.*Journal of Autism and Developmental Disorders*, *46*(7), 2327-2339.doi: 10.1007/s10803-016-2760-2.

Cohen, D. J., Volkmar, F. R., & Paul, R. (1986). Introduction: Issues in the classification of pervasive developmental disorders: History and current status of nosology. *Journal of the American Academy of Child Psychiatry*, *25*(2), 158-161. DOI: https://doi.org/10.1016/S0002-7138(09)60221-1.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry*, *45*(4), 719-726. DOI:10.1111/j.1469-7610.2004.00266.x.

Eaves, L. C., Ho, H. H., & Eaves, D. M. (1994). Subtypes of autism by cluster analysis. *Journal of Autism and Developmental Disorders*, *24*(1), 3-22.

Fein, D., Stevens, M., Dunn, M., Waterhouse, L., Allen, D., Rapin, I., & Feinstein, C. (1999). Subtypes of pervasive developmental disorder: Clinical characteristics. *Child Neuropsychology*, *5*(1), 1-23.DOI: 10.1076/chin.5.1.1.7075.

Frazier, T. W., Youngstrom, E. A., Speer, L., Embacher, R., Law, P., Constantino, J., ... & Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(1), 28-40. doi: 10.1016/j.jaac.2011.09.021.

Folstein, S. E., Bisson, E., Santangelo, S. L., &Piven, J. (1998). Finding specific genes that cause autism: a combination of approaches will be needed to maximize power. *Journal of autism and developmental disorders*, *28*(5), 439-445. doi.org/10.1023/A:102600860.

Garmezy, N. E., & Rutter, M. E. (1983). Stress, coping, and development in children. In *Seminar on Stress and Coping in Children, 1979, Ctr for Advanced Study in the Behavioural Sciences, Stanford, CA, US*: Johns Hopkins University Press.

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., ... & Smith, I. (2013). Investigating phenotypic heterogeneity in children with autism spectrum disorder: a factor mixture modeling approach. *Journal of Child Psychology and Psychiatry*, *54*(2), 206-215. doi: 10.1111/j.1469-7610.2012.02588.x.

Grzadzinski, R., Huerta, M., Lord, C. (2013).DSM-5 and autism spectrum disorders (ASDs): an opportunity for identifying ASD subtypes. *Molecular Autism, 4*(12).doi.org/10.1186/2040-2392-4-12.

Hu, V. W., & Steinberg, M. E. (2009). Novel clustering of items from the Autism Diagnostic Interview-Revised to define phenotypes within autism spectrum disorders. *Autism Research*, *2*(2), 67-77.doi: 10.1002/aur.72.

Lord, C., Rutter, M., Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*. 24 (5), 659–85. doi:10.1007/BF02172145.

Lotter, V. (1974). Social adjustment and placement of autistic children in Middlesex: A follow-up study. *Journal of Autism and Childhood Schizophrenia*, *4*(1), 11-32. doi.org/10.1007/BF02104997.

Matson, J. L., &Mulick, J. A. (1991). *Handbook of mental retardation*.new York: Pergamon Press.

Palmer, C. J., Paton, B., Enticott, P. G., &Hohwy, J. (2015). 'Subtypes' in the presentation of autistic traits in the general adult population. *Journal of Autism and Developmental Disorders*, *45*(5), 1291-1301. doi: 10.1007/s10803-014-2289-1.

Persico, A. M., &Bourgeron, T. (2006). Searching for ways out of the autism maze: genetic, epigenetic and environmental clues. *Trends in neurosciences*, *29*(7), 349-358.doi: 10.1016/j.tins.2006.05.010.

Rescorla, L. (1988). Cluster analytic identification of autistic preschoolers. *Journal of Autism and Developmental Disorders*, *18*(4), 475-492. doi.org/10.1007/BF02211868.

Ring, H., Woodbury-Smith, M., Watson, P., Wheelwright, S., & Baron-Cohen, S. (2008). Clinical heterogeneity among people with high functioning autism spectrum conditions: evidence favouring a continuous severity gradient. *Behavioural and Brain Functions*, *4*(1), 11. doi: 10.1186/1744-9081-4-11.

Sevin, J. A., Matson, J. L., Coe, D., Love, S. R., Matese, M. J., & Benavidez, D. A. (1995). Empirically derived subtypes of pervasive developmental disorders: A cluster analytic study. *Journal of Autism and Developmental Disorders*, *25*(6), 561-578.

Siegel, B., Anders, T. F., Ciaranello, R. D., Bienenstock, B., & Kraemer, H. C. (1986). Empirically derived subclassification of the autistic syndrome. *Journal of Autism and Developmental Disorders*, *16*(3), 275-293.dx.doi.org/10.1007/BF01531660

Snow, A. V., Lecavalier, L., &Houts, C. (2009). The structure of the Autism Diagnostic Interview-Revised: diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry*, *50*(6), 734-742.dx.doi.org/10.1111/j.1469-7610.2008.02018.x

Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R. M., &Risch, N. (2002). Behavioural phenotypic variation in autism multiplex families: evidence for a continuous severity gradient. *American Journal of Medical Genetics Part A*, *114*(2), 129-136.

Stevens, M. C., Fein, D. A., Dunn, M., Allen, D., Waterhouse, L. H., Feinstein, C., &Rapin, I. (2000). Subgroups of children with autism by cluster analysis: A

longitudinal examination. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(3), 346-352. DOI:10.1097/00004583-200003000-00017.

Szatmari, P., Jones, M. B., Zwaigenbaum, L., & MacLean, J. E. (1998). Genetics of autism: overview and new directions. *Journal of Autism and Developmental Disorders*, *28*(5), 351-368./dx.doi.org/10.1023/A:1026096203946.

Tsai, L. Y. (1992). Diagnostic issues in high-functioning autism. In *High-functioning individuals with autism* (pp. 11-40). Boston, MA: Springer. doi.org/10.1007/978-1-4899-2456-8_2.

Veatch, O. J., Veenstra-VanderWeele, J., Potter, M., Pericak-Vance, M. A., & Haines, J. L. (2014). Genetically meaningful phenotypic subgroups in autism spectrum disorders. *Genes, Brain and Behaviour*, *13*(3), 276-285.doi: 10.1111/gbb.12117.

Walker, D. R., Thompson, A., Zwaigenbaum, L., Goldberg, J., Bryson, S. E., Mahoney, W. J., ... &Szatmari, P. (2004). Specifying PDD-NOS: a comparison of PDD-NOS, Asperger syndrome, and autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(2), 172-180. DOI:10.1097/00004583-200402000-00012.

Waterhouse, L., Wing, L., Spitzer, R., Siegel, B. (1992). Pervasive developmental disorders: from DSM-III to DSM-III-R. *J Autism Dev Disord*., 22(4),525-49.

Wiggins, L. D., Robins, D. L., Adamson, L. B., Bakeman, R., & Henrich, C. C. (2012). Support for a dimensional view of autism spectrum disorders in toddlers. *Journal of autism and developmental disorders*, *42*(2), 191-200. DOI: 10.1007/s10803-011-1230-0

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, *9*(1), 11-29.

Witwer, A. N., &Lecavalier, L. (2008). Examining the validity of autism spectrum disorder subtypes. *Journal of Autism and Developmental Disorders*, *38*(9), 1611-1624.dx.doi.org/10.1007/s10803-008-0541-2.

Table 1

ReviewedStudieson the Classification ofAutism Spectrum Disorders

Study	Sample	Age	Diagnostic criteria	Variables	Number of clusters	Proposed Model
Eaves et al., (1994)	166	6 months-12 years 4 months	No specific instrument. Diagnosis was according DSM-III and DSM-III-R.	43 items	4	Distinct Subtypes
Sevin et al., (1995)	34	2-22 years	CARS RLRS ABC VABS AAMD	7 (based on scores on the instruments)	4	Distinct Subtypes
Fein et al., (1999)	194	36-84 months	Autism Diagnostic Checklist VABS	Nonverbal ratio IQ and Vineland subdomain scores	2	Distinct Subtypes

Stevens et al., (2000)	194	4.7 years(mean)	Autism Diagnostic Checklist SB PPVT WADIC VABS	Expressive and Receptive measures of Language, Nonverbal Intelligence and Social Behaviour	2	Distinct Subtypes
Spiker et al., (2002)	351	9.3 years (mean)	ADI-R ADOS	12 based on ADI-R scores	3	Severity Gradient
Bitsika et al., (2008)	53	4-12 years	VABS CARS ASDBC	8 dependant variables Age Verbal and full IQ Wechsler Scale VABS and CARS		Severity Gradient
Hu & Steinberg, (2009)	1,954	1.85-47.68 years	ADI-R PPVT	123 item scores	4	Support for both Distinct groups

						and the Gradient Model
Wiggins et al., (2012)	186	41-79 months	ADI-R ADOS CARS MSEL VABS	CARS		Severity Gradient
Veatch et al., (2014)	1,261 2,563 (replication sample)	2-21 years	ADI-R ADOS VABS	16 domain scores	2	Distinct Subtypes
Cholemkery et al., (2016)	463	3-21 years	ADI-R ADOS SRS Social Communication Questionnaire CBCL PPVT K-ABC	Based on ADI-R subscales and DSM-5 criteria	3	Support for both Distinct groups and the Gradient Model

CARS Childhood Autism Rating Scale, *RLRS* Ritvo-Freeman Real Life Rating Scale, *ABC* Autism Behaviour Checklist, *VABS* Vineland Adaptive Behaviour Scales, *AAMD* Adaptive Behaviour Scale, *SB* Stanford-Binet Intelligence Scale, *PPVT* Peabody Picture Vocabulary Test, *WADIC* Wing Autistic Disorders Interview Checklist, *ADI-R* Autism Diagnostic Interview-Revised, *ADOS* Autism Diagnostic Observation Schedule, *ASDBC* the Autism Spectrum Disorder Behaviour Checklist, *MSEL* Mullen Scales of Early Learning, *SRS* Social Responsiveness Scale, *CBCL* Child Behaviour Checklist, *K-ABC* Kaufman Assessment Battery for Children