

Structural equation analysis of a hypothesised symptom model in the autism spectrum

Natasja D.J. van Lang,¹ Anne Boomsma,² Sjoerd Sytema,³ Annelies A. de Bildt,¹
Dirk W. Kraijer,¹ Cees Ketelaars,¹ and Ruud B. Minderaa¹

¹University of Groningen, Department of Child and Adolescent Psychiatry, The Netherlands; ²University of Groningen, Department of Statistics & Measurement Theory, The Netherlands; ³University of Groningen, Department of Psychiatry, The Netherlands

Background: Several studies showed a different symptom structure underlying the spectrum of autistic-like disorders from the behaviour triad as mentioned in the DSM-IV. In the present study, a hypothesised symptom model for autism was constructed, based on earlier explorative findings, and was put to a confirmatory test. **Method:** Items from the Autism Diagnostic Interview-Revised (ADI-R) were used to examine the goodness of fit of the DSM-IV model, the hypothesised symptom model, and two additional models for autism. All models were tested in a group of 255 verbal and nonverbal individuals with minor to severe autistic symptomatology. **Results:** The DSM-IV model encountered estimation problems. Conversely, the hypothesised symptom model had no such problems and proved to have a better fit to the sample data than the two additional models for autism. However, some of the observed variables were weak indicators of the three latent factors in the model. **Conclusions:** The hypothesised symptom model appeared to be a plausible model in a group of individuals with a broad range of autistic behaviours and levels of functioning. Nevertheless, the stability of the model needs further examination in a larger group of individuals with disorders in the autism spectrum, and with varying degrees of intellectual functioning. **Keywords:** Autism spectrum disorder, symptom structure, structural equation modelling.

Many studies have revealed that autism is a genetically determined disorder (Rutter, Silberg, O'Connor, & Simonoff, 1999), and recent studies showed this genetic vulnerability also for the broader autism phenotype (Constantino & Todd, 2000; Silverman et al., 2002; Spiker, Lotspeich, Dimiceli, Myers, & Risch, 2002). Although autism has a neurodevelopmental origin, it is nevertheless defined by its behavioural properties. In the DSM-IV-TR (American Psychiatric Association, 2000), autism is described by impairments in three behaviour domains (the DSM triad): qualitative impairments in social interaction, qualitative impairments in communication, and the occurrence of stereotyped behaviours or restricted interests. To examine the relationship between candidate genes and autistic behaviour domains, it is important to use behaviour domains that have proved to be empirically valid. For instance, in studies that examined the three behaviour domains of the DSM triad, it was shown that these behaviour domains are highly interrelated (Spitzer & Siegel, 1990; Volkmar et al., 1994; Wing & Gould, 1979). Moreover, recent studies that dealt with the problem of whether the underlying symptom structure in autism is similar to the three behaviour domains of the DSM triad provided evidence of a different symptom structure. Factor analytic studies on autistic core features showed a symptom structure that captured the following behaviour domains: 'joint attention', 'affective reciprocity' and 'theory of mind' (Robertson, Tanguay,

L'Ecuyer, Sims, & Waltrip, 1999; Tanguay, Robertson, & Derrick, 1998), 'autistic symptomatology' and 'level of functioning' (Szatmari et al., 2002), or 'spoken language', 'social intent', 'compulsions', 'milestones', 'savant skills', and 'sensory aversions' (Tadevosyan-Leyfer et al., 2003).

In accordance with this research, two earlier explorative studies by the present authors also gave evidence of a different symptom structure in autism. In these studies (Van Lang et al., internal reports), the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994) was used as a framework to establish behaviour or symptom domains in homogeneous groups of respectively 87 verbal individuals with an autism spectrum disorder and 156 verbal individuals with an autism spectrum disorder or with mild autistic traits. The ADI-R is a standardised diagnostic interview for primary caregivers, and yields a comprehensive description of a child's behaviour over two time periods: the 4–5 age period, and the current status of functioning. A classification of an Autistic Disorder is made with an ADI-R algorithm that consists of 37 extracted items. The items that are specifically related to so-called verbal individuals were used accordingly in both studies. Depending on the applied time period, two or three candidate symptom domains appeared to have best face validity: impaired social communication, stereotyped features in speech and behaviour, and impaired play skills.

Impaired social communication was composed mostly of items from the DSM domain ‘impaired social interaction’, but also extended to items about inadequate use of gestures and failure to initiate or sustain conversational interchange from the DSM domain ‘impaired communication’. *Stereotyped features in speech and behaviour* was reflected by items of the DSM domain ‘stereotyped behaviours and restricted interests’, but it also included items about stereotyped, repetitive or idiosyncratic speech characteristics from the DSM domain ‘impaired communication’. *Impaired play skills* emerged as a separate factor when ratings of the current status of functioning were applied, combining items from the DSM domains ‘impaired social interaction’ and ‘impaired communication’, i.e., failure to develop peer relationships and lack of varied spontaneous make-believe or social imitative play.

Because these symptom domains were extracted from two explorative techniques that applied to verbal individuals only, it is necessary to investigate whether this hypothesised symptom model also holds in a larger sample of verbal and nonverbal individuals using confirmatory test procedures. Structural equation modelling was chosen as the statistical technique to test this hypothesised symptom model in a sample of $N = 255$ verbal and nonverbal individuals with minor to severe autistic symptomatology. In general, a structural equation model is a combination of a measurement model and a (structural) latent variable model (Bollen, 1989). The models considered in the present study are measurement models with correlated latent variables that represent the symptom domains. First, the DSM triad model was estimated and tested in two conditions: applying ADI-R ratings from the 4–5 age period (Model A1), and applying ADI-R ratings from the current age status (Model A2). Second, the hypothesised symptom model was estimated and tested, applying ADI-R ratings from the current age status (Model B2). In addition, the results for Model B2 were compared to those of two additional models for autism: (a) a one-factor model where all autistic behaviours were hypothesised to represent one autism factor, and (b) a two-factor model that represented a combined factor of impaired social interaction and communication, and a factor of stereotyped and restricted behaviours.

Method

Investigation of the models

Symptom domains were examined by using the 12 ADI-R labels that are described in the ADI-R algorithm. These ADI-R labels are sum-scores of three to four items that are directly related to the DSM criteria of an Autistic Disorder, reflecting the DSM triad of qualitative impairments in reciprocal social interaction (S), qualitative impairments in communication (C), and having

repetitive behaviours and stereotyped patterns (R). To investigate whether the 12 ADI-R labels could be used as indicators for the symptom domains (or latent factors) in the models tested, item factor loadings were examined first. Based on the two earlier explorative studies, it was shown that all ADI-R items had standardised factor loadings $\geq .30$. McDonald (1999, p. 177) indicated that items that are used in scales need to be sufficiently homogeneous with minimal standardised factor loadings of $.30$. It was therefore decided that the 12 ADI-R labels could be used as indicators for the symptom domains. The 12 ADI-R labels or indicators represent the following autistic behaviours: failure to use nonverbal behaviours to regulate social interaction (S1); failure to develop peer relationships (S2); lack of shared enjoyment (S3); lack of socio-emotional reciprocity (S4); lack of, or delay in, spoken language and failure to compensate through gesture (C1); lack of varied spontaneous make-believe or social imitative play (C2); relative failure to initiate or sustain conversational interchange (C3); stereotyped, repetitive or idiosyncratic speech (C4); encompassing preoccupation or circumscribed pattern of interest (R1); compulsive adherence to non-functional routines or rituals (R2); stereotyped and repetitive motor mannerisms (R3); preoccupations with part-objects or non-functional elements of materials (R4).

The first factor model to be tested (Model A1) was the DSM-IV triad with ratings applying to the 4–5 age period. These ratings correspond directly to the DSM criteria for Autistic Disorder. The following measurement model was examined ($\mathbf{x} = \Lambda\xi + \delta$), where the observed variables x_i ($i = 1, 2, \dots, 12$) are a linear function of a latent factor ξ_j ($j = 1, 2, 3$) and a measurement error δ_i . In this three-factor model, the latent variables or factors are the three DSM symptom domains: ‘impaired social interaction’ (ξ_1) with four indicators (S1 to S4), ‘impaired communication’ (ξ_2) with four indicators (C1 to C4), and ‘stereotyped behaviours’ (ξ_3) with four indicators (R1 to R4).

The second factor model (Model A2) was similar to Model A1, i.e., the structure of the DSM-IV triad, but now with ratings that applied to the current age status. It is hypothesised that ratings of the current age status reflect a more reliable picture of a child’s behaviour, especially when the individuals are amply aged above 4–5 years. In addition, such ratings are supposed to be less subject to information bias by a child’s primary caregiver.

The third factor model (Model B2) was the hypothesised symptom model based on earlier explorative findings, with ratings of subject’s current age status. This measurement model has three hypothesised symptom domains: ‘impaired social communication’ (ξ_1) with five indicators (S1, S3, S4, C1 and C3), ‘stereotyped language and behaviours’ (ξ_2) with five indicators (C4, R1, R2, R3 and R4), and ‘impaired make-believe and play skills’ (ξ_3) with two indicators (S2 and C2). It should be realised that although the last factor has only two indicators, these indicators are sum-scores of eight ADI-R items about impaired play skills.

Two additional models for autism were estimated and tested. A one-factor model was constructed by comprising all 12 ADI-R indicators into one symptom domain ‘autistic features’. In addition, a two-factor model was

constructed by combining eight indicators about impaired social interaction and communication (*S1* to *C4*) into the symptom domain 'impaired social communication', and four indicators about repetitive and stereotyped behaviours (*R1* to *R4*) into the symptom domain 'stereotyped language and behaviours'.

For all models under investigation, it was assumed that the covariances between the factors were non-zero.

Sample

From 308 participants with autistic symptomatology and with ADI-R data available, 255 children and adolescents with a full-scale intelligence quotient (FIQ) larger than 20 were selected. The participants were recruited by two different designs. First, 209 participants were recruited by a population-based screening for pervasive developmental disorders at schools for children with mild to severe learning problems. All these individuals were likely to have a disorder on the autism spectrum according to the PDD-MRS, a scale completed by teachers and specifically designed to detect pervasive developmental disorders in this population (Kraijer, 1997). Second, 46 participants were recruited by a clinical study in an Outpatient Clinic for patients with Autism Spectrum Disorders. All participants ($N = 255$) were examined in detail on their autistic behaviours with the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994) and the Autism Diagnostic Observation Schedule (ADOS-G; Lord et al., 2000), and were classified accordingly. In addition, based on the ADI-R protocols and ADOS-G videotapes, each participant received a clinical judgement of having an autism spectrum disorder or not, made by four experienced clinicians (for a detailed description of the procedure, see De Bildt et al., 2004). The cognitive abilities of each participant were determined either by an established FIQ if it had been determined within two years of the study, or by an reassessed FIQ with Dutch versions of the WAIS-R (Uterwijk, 2000), WISC-R (Van der Steene et al., 1986), WPPSI-R (Van der Steene & Bos, 1997) or the Dutch nonverbal intelligence scale SON-R (Snijders & Snijders-Oomen, 1975). Participants with profound intellectual disability ($FIQ \leq 20$) and those who were not testable in a standard test situation were excluded from the analyses.

In this group of 255 participants, 130 received a clinical judgement of a disorder on the autism spectrum (ASD) according to the DSM-IV-TR criteria. The other 125 were judged to have social or communication problems, but their behaviours were not severe enough to warrant a clinical diagnosis of an autism spectrum disorder. These participants were defined as 'typically developing subjects'. The ASD group was composed of 54 participants with an Autistic Disorder, 3 with Asperger's Syndrome, and 73 with a Pervasive Developmental Disorder-Not Otherwise Specified. The majority of the participants were male: 196 boys and 59 girls. The chronological ages varied between 4 and 20 years, with a mean of 11;03 (years and months), and a standard deviation of 3;11. The FIQ ranges varied between 20 and 129, with most participants falling in the severe ($N = 66$), moderate ($N = 52$), and mild ($N = 88$) FIQ range, and relatively less in the borderline ($N = 28$), normal ($N = 15$), or above normal ($N = 6$) FIQ range. In

Table 1 Sample characteristics: Mean scores (SD) of age and FIQ, and on the ADI-R domains 'impaired social interaction', 'impaired communication', and 'stereotyped behaviour'

	TD	PDD-NOS/AS	AD
N	125	76	54
Age (years; months)	11;06	12;01	11;03
FIQ*	49 (19)	61 (27)	55 (24)
ADI-R domain 'impaired social interaction'	12.0 (7.2)	19.0 (5.7)	21.1 (6.0)
ADI-R domain 'impaired communication'	8.6 (4.9)	13.9 (4.6)	14.2 (4.7)
ADI-R domain 'stereotyped behaviour'	2.7 (2.6)	5.0 (2.6)	6.3 (2.5)

Note: TD means 'typically developing subjects'; PDD-NOS/AS denotes a combination of Pervasive Developmental Disorder-Not otherwise Specified and Asperger's Syndrome; AD means Autistic Disorder.

*A significant main effect of group with $p < .001$.

Table 1, sample characteristics are presented for the 'typically developing subjects' (TD), for the combination of participants with a Pervasive Developmental Disorder-Not Otherwise Specified or Asperger's Syndrome (PDD-NOS/AS), and those with an Autistic Disorder (AD).

An analysis of variance (ANOVA) showed a significant main effect of group on FIQ, with post hoc analyses indicating that the TD group had a significantly lower FIQ than the PDD-NOS/AS group ($p < .01$). Multivariate analysis of variance (MANOVA) showed a significant main effect of group on the three ADI-R domain scores, and post hoc analyses revealed that the TD group had significantly lower scores on the three ADI-R domains, compared to the PDD-NOS/AS group and the AD group ($p < .01$).

Model estimation

A Maximum Likelihood (ML) estimation procedure was used to examine the goodness of fit of the models. The assumptions of the ML procedure are: (1) the sample observations are independently distributed, (2) the indicators, or observed variables, have a multivariate normal distribution, (3) the hypothesised model is approximately correct, (4) a sample covariance matrix \mathbf{S} is being analysed, and (5) a large sample size N is used for a proper approximation of asymptotic properties of parameter, standard error and model-fit estimators (Jöreskog & Sörbom, 1989). With respect to the requirements of multivariate normal distribution, the 12 ADI-R indicators were not normally distributed (a median interquartile range of 2.00, a median skewness of .95, with minimum .14 and maximum 1.40, and a median kurtosis of .01, with minimum -1.08 and maximum 1.14). Although the skewness and kurtosis values were not extreme for maximum likelihood estimation (cf. Boomsma & Hoogland, 2001), it was decided to use a robust ML estimation procedure to improve on the estimates of standard errors and model fit while analysing a sample covariance matrix \mathbf{S} . The global model fit was evaluated using the scaled, i.e., mean-adjusted, chi-square statistic of Satorra and Bentler (1994). This robust ML estimation procedure

can be summarised as follows. First, the sample covariance matrix \mathbf{S} , and the corresponding estimate of the asymptotic covariance matrix of the sample covariances, $AC\hat{ov}(\mathbf{S})$, were calculated using the PRELIS program 2.54. Second, the estimated covariance matrix \mathbf{S} and $AC\hat{ov}(\mathbf{S})$ were used as input for the LISREL program 8.54 to estimate the postulated models using robust ML estimation. The PRELIS 2.54 program is part of the LISREL program 8.54 (Jöreskog & Sörbom, 1996).

Results

Models A1 and A2

Both DSM models (Model A1 and Model A2) encountered an estimation problem that could be labelled as an empirical identification problem. It appeared that the covariance matrix of the latent factors in both models was not positive definite, showing a correlation larger than one between the factors S (DSM domain of impaired social interaction) and C (DSM domain of impaired communication). These findings suggest a high multicollinearity between the factors S and C when ADI-R ratings of a child's age of 4–5 years and ADI-R ratings of subject's current functioning were applied.

To examine whether the cause of the problem could be the random sample data, or the postulated model, or both, the nature of the estimation problem was further investigated. Two random sub-samples of sizes $N = 205$ and $N = 155$ were taken from the total sample size of $N = 255$ to investigate whether the problem might be due to the sample covariance matrix \mathbf{S} . For each of these two sub-samples, Models A1 and A2 were estimated, and in all cases the problem of an improper covariance matrix of the latent variables remained. In addition, different starting values were used, and a different estimation

program (EQS 5.4; Bentler, 1995) was employed as an additional check on the results. In both cases, however, the problem remained.

Furthermore, Models A1 and A2 were tested with the restriction of zero covariances between the factors (i.e., factors were assumed to be uncorrelated). The results showed inadequate fit measures, and modification indices pointing to a high correlation between S and C . In addition, Models A1 and A2 were tested with the restriction that the factor correlation matrix is strictly positive definite, by constraining the smallest eigenvalue of the covariance matrix of the latent variables to be greater than zero (i.e., values of the parameters were restricted to border values which made the covariance matrix of the latent factors just positive definite). Apart from the fact that such forced model restrictions are most often unsatisfactory, i.e., in our case an almost perfect high estimated correlation between the two factors, the results of both restricted models still showed an inadequate fit (see the goodness-of-fit values for the restricted Model A2 in Table 2). Therefore, it was concluded that the irregularities encountered were (primarily) due to the postulated model: Models A1 and A2 cannot be properly estimated and are therefore implausible. Subsequently, it makes no sense to discuss the size of parameter estimates and corresponding standard errors of these two models. For a detailed presentation of the LISREL output regarding the estimation of Models A and B, the reader is referred to the Internet site: <http://www.ppsw.rug.nl/~boomsma/lang.htm>.

Model B2

In contrast to Models A1 and A2, no estimation problems were encountered with Model B2.

Table 2 Goodness-of-fit values for Model B2, for the one- and the two-factor model for autism, and for the restricted Model A2. The cut-off criteria mentioned in the notes are partly based on the findings of Hu and Bentler (1999)

	Global fit measures ¹						RMSEA ² (90% C.I.)	SRMR ³	NNFI ⁴	AIC ⁵	ECVI ⁶
	<i>df</i>	χ^2_{NWS}	<i>p</i>	χ^2_{SB}	<i>p</i>						
Model B2	51	74.20	.02	65.59	.08	.03 (.00–.06)	.05	.98	119.59	.47	
One-factor model	54	375.81	.00	324.36	.00	.14 (.13–.16)	.10	.76	372.36	1.47	
Two-factor model	53	308.90	.00	268.73	.00	.13 (.11–.14)	.10	.79	318.73	1.25	
Model A2 (restricted)	51	302.74	.00	264.14	.00	.13 (.11–.14)	.10	.72	318.14	1.25	

¹ The global fit measures are the Normal Theory Weighted Least Squares chi-square (χ^2_{NWS}) and the mean-adjusted chi-square of Satorra–Bentler (χ^2_{SB}), taking the non-normality of the data into account. Models with a scaled χ^2 -value with $p > .05$ are judged to have a reasonable model fit to the observed data.

² The Root Mean Square Error of Approximation (RMSEA) is the value of the model test statistic for close fit. Values $\leq .06$ may indicate a reasonable fit. C.I. is the 90% confidence interval.

³ The Standardised Root Mean Square Residual (SRMR) is a standardised measure of the average of fitted residuals. Values $\leq .08$ may indicate a reasonable fit.

⁴ The Nonnormed Fit Index (NNFI) is a measure of the improved fit of the postulated model relative to the independence model. Values $\geq .95$ may indicate a reasonable fit.

⁵ The Akaike's Information Criterion (AIC) is a fit index that takes the model complexity into account. This index can be used to compare the relative fit of different models, with lower values indicating a better fit.

⁶ The Estimated Cross Validation Index (ECVI) gives an estimate of the stability of parameter estimates in future samples. This index can be used to compare fit over models, with lower values indicating a better predictive fit.

Table 3 Estimated correlation matrix of the latent factors of Model B2. Presented are the estimated correlations, the corresponding estimated standard errors (in parentheses), and the test statistics (*t*-values) for each latent factor

	1. Impaired social communication	2. Impaired play skills	3. Stereotyped language/behaviour
1.			
2.	.48, (.06), 8.09		
3.	.48, (.07), 6.98	.37, (.07), 5.27	

The overall fit of model B2 was found to be adequate: the Satorra–Bentler chi-square statistic was 65.59, with 51 degrees of freedom, and a *p*-value of .8, and the ML cut-off criteria suggested by Hu and Bentler (1999) were met for the estimated values of RMSEA, SRMR and NNFI. The estimated correlation matrix of the latent factors appeared to be proper (see Table 3). The parameter estimates could be identified and revealed that the correlations between the three latent factors were substantively different from zero. In addition, low standard errors of all estimated model parameters were found that indicated an acceptable stability (reliability) of these estimates. The test statistics (*t*-values) for the null hypothesis of population parameters having a value of zero were all larger than five.

The completely standardised estimates of the factor loadings are presented in Figure 1.

The completely standardised estimates of the factor loadings vary between .47 and .99, and R^2 , squared multiple correlation coefficient for the indicators, varies between .22 and .97 (estimates not shown in this paper). However, some indicators appear to have a relatively weak association with their corresponding factor. This may partly be explained on the basis of the frequency distribution of the responses of the participants on some of the items. The majority of the participants had a score of zero on the items about stereotyped features in speech and behaviours (*R1* to *R4*), or on the item about 'impaired initiation and maintenance of conversational interchange' (*C3*). In general, these items represent behaviours that are most prominent for individuals with AD.

In addition, most subjects had intelligence scores that fell in the severe to borderline FIQ range. To investigate whether Model B2 would hold in a group of only individuals with an intellectual disability, Model B2 was tested in a sub-sample of individuals with a FIQ score < 85 ($N = 234$). Compared to the results of the total sample, no substantial differences in model fit were found: e.g., $df = 51$; $\chi^2_{NWLS} = 79.82$ with $p = .01$; $\chi^2_{SB} = 69.94$ with $p = .04$; RMSEA = .04 (90% C.I.: .00–.19), SRMR = .06, and NNFI = .98. Although the two samples differed only by 24 individuals, it can be noticed that the fit in the smaller sample was slightly worse (notably the range of the confidence interval for the RMSEA). Clearly, Model B2 needs to be further validated in an independent

sample to examine whether it would also hold for individuals with FIQ scores larger than 85.

Comparing Model B2 with two additional factor models for autism

Table 2 shows that the fit of the one- and the two-factor model for autism is highly inadequate. In comparing all models under study, the values of Akaike's information criterion (AIC) and the expected cross-validation index (ECVI) show that Model B2 can be trusted to have comparatively well-fitting properties.

Discussion

The hypothesised symptom model with three latent factors or symptom domains fitted the sample data reasonably well. Both DSM models (with ratings applying to the 4–5 age period and to the current age status of the child) revealed estimation problems that indicated a high correlation between the DSM domains 'impaired social interaction' and 'impaired communication'. In contrast, Model B2 did not encounter such estimation problems. Model B2 is based upon a combined construct of impairments in social interaction and communicating skills, and showed no substantive discrepancies between the sample data and model-implied covariances. In addition, Model B2 fitted the sample data better than the two additional factor models for autism: the one-factor model, and the correlated two-factor model which consisted of a factor for the combined construct of impairments in social interaction and communication, and a factor for repetitive and stereotyped behaviours. The results clearly showed a substantive lack of fit of these two additional models, unlike Model B2. Therefore, Model B2 might provide a useful measurement structure for future studies that aim to explore the relationship between symptom domains and, for instance, autistic susceptibility genes.

The three latent factors in Model B2 are based on the symptomatology of autism, and have a different item structure than that of the DSM triad. In Model B2, *Impaired social communication* contains information about poor verbal and nonverbal social communicative interchange. *Impaired make-believe and play* is comprised of a lack of play skills in individual activities and in relationship with peers. *Stereotyped language and behaviour* consists of stereotyped characteristics in speech and behaviour. These three latent factors were identified in two earlier studies in which explorative techniques were used (Van Lang et al., internal reports). They emerged as underlying constructs for 87 verbal individuals with an established disorder on the autism spectrum, and for 156 verbal individuals with a broader autism phenotype, all with FIQ scores

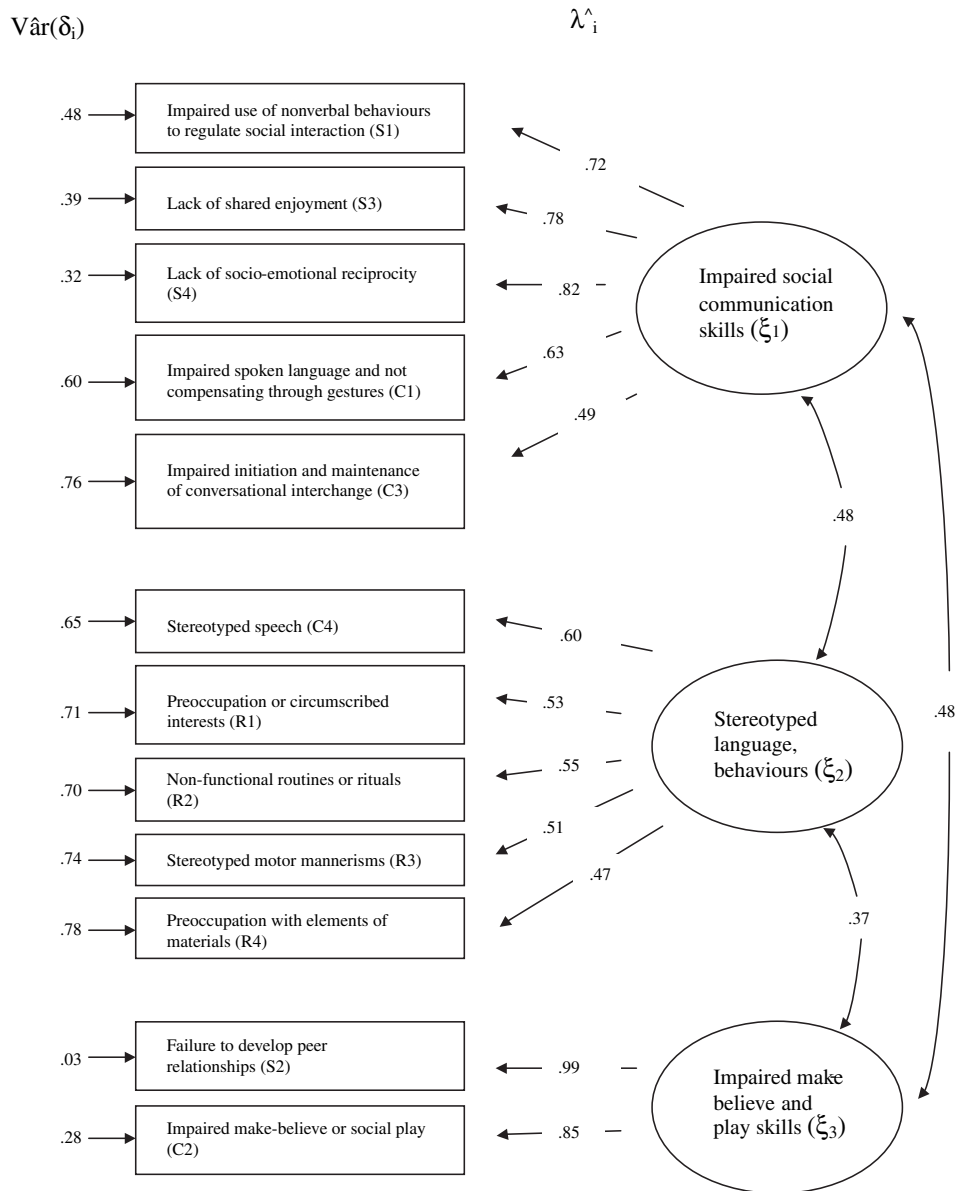


Figure 1 Confirmatory analysis on the hypothesised Model B2. Presented are the completely standardised estimates of factor loadings ($\hat{\lambda}_i$) of the indicators, the estimated correlations between the symptom domains ζ , and the estimated variances of the measurement errors, $V\hat{a}r(\delta_i)$

larger than 35. In the present study, the sample was expanded to 255 verbal and nonverbal individuals with a broad scope of various autistic behaviours and with a FIQ score larger than 20. The results showed that the validity of the latent variable structure from the explorative studies was supported in this confirmatory study.

Evidence for a different symptom structure than the DSM triad emerged from other studies as well (Robertson et al., 1999; Szatmari et al., 2002; Tanguay et al., 1998; Tadevosyan-Leyfer et al., 2003). These authors also used the ADI-R as the framework for investigating the behaviour structure in autism. Conditional upon the included sample and the applied ADI-R indicators (item scores or domain scores), the autism symptomatology was categorised in two or

three separate symptom domains or it was combined with a distinct domain of adaptive functioning. When the latent factors of Model B2 are compared with the results of the group of Robertson and Tanguay (Robertson et al., 1999; Tanguay et al., 1998), it turns out that the factor ‘impaired play skills’ of Model B2 has similarities with their factor ‘theory of mind’. In addition, the factor ‘impaired social communication’ of Model B2 may be a combination of their factors ‘joint attention’ and ‘affective reciprocity’. However, the factor ‘stereotyped language/behaviours’ of Model B2 cannot be compared, because stereotypes in behaviour were not included in the analyses made by the group of Robertson and Tanguay.

With regard to the first three factors found by Tadevosyan-Leyfer and colleagues (2003), the factor

'impaired social communication' of Model B2 resembles their factor 'social intent', which again underlines the high overlap between the DSM domains 'impaired social interaction' and 'impaired communication'. However, the factor 'stereotyped language/behaviours' of Model B2 was not found by Tadevosyan-Leyfer and her colleagues, although their factors 'spoken language' and 'compulsions' seem to represent behaviours that were included in the factor 'stereotyped language/behaviours' of Model B2.

In a review study by Beglinger and Smith (2001), it was concluded that there is no system yet available that accounts for the symptom heterogeneity in autism. Based on their review, the authors proposed a dimensional conceptualisation for autism, in which four domains are identified: variations in developmental delays, in social impairments, in restricted behavioural features, and in FIQ. The factors from Model B2 give partial support for this proposed conceptualisation of autism, but also reveal an additional domain of impairments in play skills.

Nevertheless, it is essential to investigate the validity of Model B2 in a new sample of individuals with an autism spectrum disorder or with autistic traits, since the present study has a number of limitations that need to be addressed. For instance, the present study was limited by the fact that Model B2 was constructed by results obtained from two explorative studies that included selections of the sample used in the present study. Given the assumptions of the (robust) maximum likelihood estimation procedure (asymptotic theory), it was decided to use as large a sample size as possible for the confirmatory analyses. However, it restricted the possibility to examine the validity of Model B2 in a completely independent sample from those being used in the two explorative studies. An independent cross-validation for Model B2 is therefore needed.

In addition, the present study included many individuals who had an intellectual disability. Although it was shown that the fit of Model B2 did not substantially change when only individuals with FIQ scores below 85 were selected, the validity of Model B2 needs to be investigated in a new sample of higher-functioning autistic individuals. Constantino and Todd (2003), for instance, found that social deficits that are characteristic for autism spectrum disorders are common and continuously distributed in a general population sample. To investigate whether not only the latent factor 'impaired social communication' but also the latent factors 'impaired make-believe and play' and 'stereotyped language and behaviours' of Model B2 will remain stable in higher-functioning autistic individuals as well, replication studies are necessary.

Finally, the present study included only 54 individuals with an Autistic Disorder. To examine

whether the latent factor 'stereotyped language and behaviour' of Model B2 is specific for these individuals, Model B2 needs to be tested in preferably larger samples of individuals with severe autistic symptoms than was possible in the present study.

However, given the heterogeneity of the sample that was used, it might be concluded that Model B2 is expected to be a fairly stable model, and that it offers a better representation of the symptom structure in autism than the DSM triad model.

Acknowledgements

This research was supported by grants from the Netherlands Organisation for Health Research and Development (ZON/MW) and from the Korczak Foundation for autism and related disorders.

Correspondence to

A. Boomsma, University of Groningen, Department of Statistics and Measurement Theory, Boumangebouw, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands; Tel: + 31 50 3636187; Fax: + 31 50 3636304; Email: A.Boomsma@ppsw.rug.nl

References

- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders*, Ed. 4 Text Revision (DSM-IV-TR). Washington, DC: American Psychiatric Association.
- Beglinger, L.J., & Smith, T.H. (2001). A review of subtyping in autism and proposed dimensional classification model. *Journal of Autism and Developmental Disorders*, 31, 411–422.
- Bentler, P.M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Boomsma, A., & Hoogland, J.J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future. A festschrift in honor of Karl Jöreskog* (pp. 139–168). Chigago, IL: Scientific Software International.
- Constantino, J.N., & Todd, R.D. (2000). Genetic structure of reciprocal social behavior. *American Journal of Psychiatry*, 157, 2043–2045.
- Constantino, J.N., & Todd, R.D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry*, 60, 524–530.
- De Bildt, A., Sytma, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., & Minderaa, R. (2004). Interrelationship between Autism Diagnostic Observation Schedule-Generic (ADOS-G), Autism Diagnostic Interview-Revised (ADI-R), and the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) classification in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, 34, 129–137.

- Hu, L., & Bentler, P.M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Jöreskog, K.G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications* (2nd edn). Chicago, IL: SPSS.
- Jöreskog, K.G., & Sörbom, D. (1996). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago, IL: Scientific Software International.
- Kraijer, D.W. (1997). *Autism and autistic-like conditions in mental retardation*. Lisse, The Netherlands: Swets & Zeitlinger.
- 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. *Journal of Autism and Developmental Disorders*, 24, 659–685.
- Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Jr, Leventhal, B.L., Dilavore, P.C., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205–223.
- McDonald, R.P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- Robertson, J.M., Tanguay, P.E., L'ecuyer, S., Sims, A., & Waltrip, C. (1999). Domains of social communication handicap in autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 738–745.
- Rutter, M., Silberg, J., O'Connor, T., & Simonoff, E. (1999). Genetics and child psychiatry: II Empirical research findings. *Journal of Child Psychology and Psychiatry*, 40, 19–55.
- Satorra, A., & Bentler, P.M. (1994). Correlations to test statistics and standard errors in covariance structure analysis. In A. von Eye & C.C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.
- Silverman, J.M., Smith, C.J., Schmeidler, J., Hollander, E., Lawlor, B.A., Fitzgerald, M., Buxbaum, J.D., Delaney, K., Galvin, P., & Autism Genetic Research Exchange Consortium (2002). Symptom domains in autism and related conditions: Evidence for familiarity. *American Journal of Medical Genetics*, 8, 64–73.
- Snijders, J.Th., & Snijders-Oomen, N. (1975). *Snijders-Oomen niet-verbale intelligentieschaal. SON 2,5–7*. Groningen, The Netherlands: Wolters-Tjeenk Willink.
- Spiker, D., Lotspeich, L.J., Dimiceli, S., Myers, R.M., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: Evidence for a continuous severity gradient. *American Journal of Medical Genetics*, 8, 129–136.
- Spitzer, R.L., & Siegel, B. (1990). The DSM-III-R field trial of pervasive developmental disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 855–862.
- Szatmari, P., Merette, C., Bryson, S., Thivierge, J., Roy, M.A., Cayer, M., & Maziade, M. (2002). Quantifying dimensions in autism: A factor-analytic study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 467–474.
- Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, P., McGrath, L., Tager-Flusberg, H., & Folstein, S. (2003). A principal components analysis of the Autism Diagnostic Interview-Revised. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42, 864–872.
- Tanguay, P.E., Robertson, J., & Derrick, A. (1998). A dimensional classification of autism spectrum disorder by social communication domains. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37, 271–277.
- Uterwijk, J.M.R. (2000). *WAIS-III: Nederlandstalige bewerking: Wechsler Adult Intelligence Scale – Ed.3: Afname en scoringshandleiding (WAIS-III 2000)*. Lisse, The Netherlands: Swets Test Publishers.
- Van Lang, N., Sytema, S., De Bildt, A., Kraijer, D., Ketelaars, C., & Minderaa, R. (internal report). *Symptom dimensions in verbal individuals with autism: A triad or a dyad?* University of Groningen: Department of Child and Adolescent Psychiatry.
- Van Lang, N., Sytema, S., De Bildt, A., Kraijer, D., Ketelaars, C., & Minderaa, R. (internal report). *Symptom structure underlying the ADOS-G and ADI-R in verbal children and adolescents with an autism spectrum disorder*. University of Groningen: Department of Child and Adolescent Psychiatry.
- Van der Steene, G., & Bos, A. (1997). *Wechsler Pre-school and Primary Scale of Intelligence-Revised. Vlaams-Nederlandse Aanpassing*. Lisse, The Netherlands: Swets & Zeitlinger.
- Van der Steene, G., Haasen, P.P., De Bruyn, E.E.J., Coetsier, P., Pijl, Y.J., Poortinga, Y.H., Lutje Spelberg, H.C., Spoelders-Claes, R., & Stinissen, J. (1986). *Wechsler Intelligence Scale for Children-Revised. Nederlandse Uitgave (WISC-RN)*. Lisse, The Netherlands: Swets & Zeitlinger.
- Volkmar, F.R., Klin, A., Siegel, B., Szatmari, P., Lord, C., Campbell, M., Freeman, B.J., Cicchetti, D.V., Rutter, M., & Kline, W. (1994). Field trial for autistic disorder in DSM-IV. *American Journal of Psychiatry*, 151, 1361–1367.
- 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, 11–29.