



Nivard, M. G., Lubke, G. H., Dolan, C. V., Evans, D., St Pourcain, M. U. B., Munafo, M., & Middeldorp, C. M. (2017). Joint developmental trajectories of internalizing and externalizing disorders between childhood and adolescence. *Development and Psychopathology*, 29(3), 919-928. <https://doi.org/10.1017/S0954579416000572>

Peer reviewed version

Link to published version (if available):
[10.1017/S0954579416000572](https://doi.org/10.1017/S0954579416000572)

[Link to publication record on the Bristol Research Portal](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via Cambridge University Press at <http://dx.doi.org/10.1017/S0954579416000572>. Please refer to any applicable terms of use of the publisher.

University of Bristol – Bristol Research Portal

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/>

Joint developmental trajectories of internalizing and externalizing disorders between childhood and adolescence.

M.G. Nivard^{1,2}, G.H. Lubke^{1,3}, C.V. Dolan¹, D.M. Evans^{4,5,6}, B. St Pourcain^{6,7}, M.R. Munafò^{6,8}, C.M. Middeldorp^{1,2,9}.

1. Biological Psychology, VU University, Amsterdam, The Netherlands
2. Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands
3. Department of psychology, University of Notre Dame, Notre Dame, USA
4. University of Queensland Diamantina institute, Translational Research Institute, Brisbane, Australia
5. School of Social and Community Medicine, University of Bristol, Bristol, UK
6. MRC Integrative Epidemiology Unit, University of Bristol, UK
7. Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
8. UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, Bristol, UK
9. Department of Child and Adolescent Psychiatry, GGZ inGeest / VU University Medical Centre, Amsterdam, The Netherlands

Acknowledgements: We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and the Wellcome Trust (Grant refs: 092731 and 102215/2/13/2) and the University of Bristol provide core support for ALSPAC This publication is the work of the authors and Michel G. Nivard will serve as guarantors for the contents of this paper. This project was supported by “Genetic influences on stability and change in psychopathology from childhood to young adulthood” (ZonMW 912-10-020) and “Genetics of Mental Illness” (ERC-230374). MGN is supported by Royal Netherlands Academy of Science Professor Award (PAH/6635). GHL is supported by ACTION. ACTION receives funding from the European Union Seventh Framework Program (FP7/2007–2013) under grant agreement No. 602768.

* Correspondence to: Michel G Nivard, m.g.nivard@vu.nl van der Boechorststraat 1, 1081BT, Amsterdam, The Netherlands

Short title: Internalizing & Externalizing symptom trajectories

Joint developmental trajectories of internalizing and externalizing disorders between childhood and adolescence.**Abstract**

This study aimed to identify trajectories of DSM-IV based internalizing (INT) and externalizing (EXT) problem scores across childhood and adolescence, and to provide insight into the co-morbidity by modeling the co-occurrence of internalizing and externalizing trajectories. INT and EXT was measured repeatedly between age 7 and age 15 years in over 7000 children and analyzed using growth mixture models. For both INT and EXT, five trajectories were identified including a very low, low, decreasing and increasing trajectory. In addition, an adolescent onset trajectory was identified for INT, and a stable high trajectory was identified for EXT. Multinomial regression showed that similar EXT and INT trajectories were associated. However, the adolescent onset INT trajectory was independent of high EXT trajectories, and persisting EXT was mainly associated with decreasing INT. Sex and early life environmental risk factors predicted EXT and, to a lesser extent, INT trajectories. The association between trajectories indicates the need to consider co-morbidity when a child presents with internalizing or externalizing disorders particularly when symptoms start early. This is less necessary when internalizing symptoms start at adolescence. Future studies should investigate the etiology of co-occurring INT and EXT, and the specific treatment needs of these severely affected children.

Keywords: internalizing disorders, externalizing disorders, developmental trajectories, comorbidity, childhood and adolescence.

Introduction

Longitudinal epidemiological studies have shown that the prevalence of joint internalizing disorders (anxiety and depressive disorders) increases from childhood to adolescence into adulthood, whereas the prevalence of externalizing disorders (Attention Deficit Hyperactivity Disorder (ADHD), conduct disorder (CD) and oppositional defiant disorder(ODD)) decreases (Costello et al., 2011). Research into the stability has shown that symptoms and disorders persist in varying degrees from childhood into adulthood (Hofstra, Van der Ende, & Verhulst, 2002; Hofstra, Van der Ende, & Verhulst, 2000). Moreover, comorbidity is common (Angold, Costello & Erkanli; 1999), not only concurrently, but also successively, as has been established in longitudinal studies: internalizing disorder in childhood predict externalizing disorder in adulthood and vice versa (Copeland, Shanahan, Costello, & Angold, 2009 ; Hofstra et al., 2000; Hofstra et al., 2002). It is well known that co-morbidity is associated with poorer outcomes (Newman, Moffitt, Caspi, & Silva (1998)).

It is important to distinguish children who are at risk for developing chronic and/or co-morbid symptoms and children whose symptoms are transient so that treatment programs can be developed that specifically target children at risk of long term psychopathology or children whose symptoms are transient. Therefore, it is necessary to investigate the existence of subgroups (classes) with distinct developmental trajectories using growth mixture modeling (GMM). Few studies using GMM have analyzed data on internalizing and externalizing problems in the period from childhood to adolescence in population based cohorts (Dekker, Ferdinand, Van Lang, Bongers, Van der Ende, & Verhulst, 2007; Letcher, Smart, Sanson, & Toumbourou, 2009; Toumbourou, Williams, Letcher, Sanson, & Smart, 2011; Van Lier, Der Ende, Koot, & Verhulst, 2007; Haltigan, Roisman, Susman, Barnett-Walker, & Monahan, 2011; Larsson, Dilshad, Lichtenstein, & Barker, 2011). No study, to the best of our knowledge, has investigated comorbidity in trajectories from childhood into adolescence. As pointed out by Angold, Costello & Erkanli (1999), research in population based cohorts is

necessary to get unbiased estimates of comorbidity and its risk factors. And since the largest changes in prevalence rates are observed in the transition from childhood to adolescence, it is critical to investigate this period.

In the current study we used GMM to model the development of DSM-IV based internalizing and externalizing problem scores (INT and EXT) measured at four occasions between age 7 and 15 years in a birth cohort of over 7,000 children. INT comprised anxiety disorders and depression, and EXT comprised ADHD, ODD and conduct disorder. The use of internalizing and externalizing summary scores is consistent with the results of several factor analytic studies of these disorders, which revealed the presence of internalizing and externalizing higher order factors explaining the covariance between individual disorders (Angold et al., 1999; Beauchaine & McNulty, 2013; Cosgrove, Rhee, Gelhorn, Boeldt, Corley, Ehringer, Young, & Hewitt, 2011). Following an initial separate analysis of internalizing and externalizing trajectories, we focused on the co-occurrence of these trajectories in a combined model. We further added well known risk factors for internalizing and externalizing psychopathology to the model as predictors of class membership such as sex, birth weight, maternal smoking during pregnancy and social class (Groen-Blokhuis, Middeldorp, van Beijsterveldt & Boomsma 2011; Weissman, Warner, Wickramaratne & Kandel 1999; Linnet, Dalsgaard, Obel, Wisborg, Henriksen, Rodriguez et al. 2003; Hack, Youngstrom, Cartar, Schluchter, Taylor, Flannery et al. 2004; Costello, Compton, Keeler, & Angold, 2003; Dolan et al 2015). An added benefit of modeling early childhood risk factors as predictors of later trajectories in a longitudinal study is that the results are not affected by participant dropout associated with these childhood risk factors (Little & Rubin, 2014). The results provide insight into the trajectories of clinically relevant internalizing and externalizing problems across childhood and adolescence as well as into the association between the internalizing and externalizing trajectories in this period.

Based on the results of previous trajectory analyses of internalizing and externalizing psychopathology measured during childhood and adolescence in population based cohorts (Dekker et al., 2007; Haltigan et al., 2011; Larsson et al., 2011; Letcher et al., 2009; Toumbourou et al., 2011; Van Lier et al., 2007) we expected for both INT and EXT a class of unaffected individuals. In addition, we expected at least a class with increasing symptoms for INT and a class with stable high and a class with decreasing symptoms for EXT. Although results of previous studies using GMM on internalizing symptoms are mixed regarding a class with persisting symptoms over time, we expected such a class given that other longitudinal studies suggest continuity over age (Hofstra et al., 2002; Hofstra et al., 2000). To the best of our knowledge, the current study is the first to investigate the co-occurrence of trajectories of internalizing and externalizing problems from childhood to adolescence. Three previous studies investigated the concordance between internalizing and externalizing trajectories in children up to age 12 (Brezo, Barker, Paris, Hébert, Vitaro, Tremblay, & Turecki, 2008; Fanti & Henrich, 2010; Fanti & Henrich, 2010; Wiggins, Mitchell, Hyde & Monk (2015)). These studies suggested that children assigned to trajectories with high scores on internalizing problems were significantly more often assigned to trajectories with moderate or high scores for externalizing problems, and vice versa. This signifies that the course of internalizing and externalizing symptoms is also associated during childhood. We expect that association to continue into adolescence.

Methods

Subjects. The Avon Longitudinal Study of Parents and Children (ALSPAC, also known as “Children of the 90s”, <http://www.bristol.ac.uk/alspac/>) is a long-term health research project (Boyd et al. 2012). More than 14,000 mothers from the Avon County in the UK were enrolled during pregnancy in 1991 and 1992, and returned at least one questionnaire. When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally.

The (psychological) health and development of these children has been followed in great detail. At ages 7, 10, 13, and 15 years, DSM-IV psychiatric disorders were assessed as part of the regular assessments. In total, 7202 children were assessed at least once for psychiatric disorders and had data available on risk factors (see table 1). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Please note that the study website contains details of all the data available through a fully searchable data dictionary (<http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary>).

Instruments. The Development And Wellbeing Assessment (DAWBA) is an instrument developed to diagnose DSM-IV psychiatric disorders (Goodman, Ford, Richards, Gatward, & Meltzer, 2000). In addition to a dichotomous variable indicating whether or not a person satisfies the criteria for a diagnosis, the DAWBA instrument can be used to calculate an ordinal “DAWBA band” score. Each band indicates the probability to suffer from a psychiatric disorder as derived from the DAWBA psychiatric interview. of the integer scores from 0 to 5 correspond to probabilities of <0.01%, 0.5%, 3%, 15%, 50%, and >70%, respectively, of satisfying the diagnostic. We analyzed the combined externalizing DAWBA band score, which includes ODD, CD, and ADHD, and the combined internalizing DAWBA band score, which includes major depression, generalized anxiety disorder, specific phobia, social phobia (at age 7, 10, 13 and 15), separation anxiety disorder (at age 7, 10 and 13) and panic disorder and agoraphobia (at age 15). All disorders were assessed by the child’s mother except for self reports of internalizing disorders at age 15.

The DAWBA band scores have shown a positive association with clinician-rated diagnosis (chance corrected kappa 0.4 to 0.7, sensitivity 0.4 to 0.8, and specificity 0.98 to 0.99), and a strong relation with indicators of mental health (Goodman, Heiervang, Collishaw, & Goodman, 2011). The internalizing DAWBA band scores (INT) and externalizing DAWBA band scores (EXT) reflect the probability of satisfying the diagnostic criteria of any internalizing or any externalizing disorder

(Goodman et al., 2011). Since category 0 did not occur in all assessments, 0 and 1 scores were collapsed into a single category (i.e., <.5%). We used the DAWBA band scores as they provide more information than the dichotomous affected/unaffected variable.

We included maternal smoking during pregnancy (no/yes), maternal highest education (6 categories), maternal and paternal social class (6 categories), maternal age at delivery, birth weight, and sex as predictors of class membership. Maternal smoking was assessed at week 18, social class and educational attainment at week 32. Maternal age at delivery and birth weight were part of the pregnancy and child baseline data.

Statistical methods. As a baseline model, we fitted a latent growth curve model to the repeated measures of INT and EXT. This model described a single trajectory that can randomly vary over individuals and included three factors, an intercept (I), linear slope (S), and quadratic slope (Q) factor, where the quadratic slope factor allows for curvilinear development. As we estimated the means (fixed effects) and the variances (random effects) of the I, S, and Q factors, this is a random effects model, implying that each child was characterized by his or her own unique growth curve (Singer & Willett, 2003).

Growth mixture models (GMM) extend the standard growth model with a latent class variable, each describing a subset of the entire population and featuring a distinct growth model within each latent class. As class membership is unknown, subjects with similar trajectories are grouped into classes in a data-driven fashion. Fixing the variances of the I, S, and Q factors to zero within each class results in a restrictive GMM (also known as latent class growth models, LCGMs), in which only average within class trajectories are estimated (i.e., means of I, S, and Q), and all variability within classes is considered to be occasion specific (Nagin, 1999). In addition to LCGMs, we also fitted random intercept models in which the variance of the intercept was not fixed to zero, thus

allowing for within class individual differences in the intercepts (Muthén & Muthén 2000; Jung & Wickrama 2008). In fitting both LGCMs and the random intercept models, we considered models with increasing number of classes. Mixture models with random slope or random quadratic terms often failed to converge and are therefore not considered.

Based on the best fitting separate models, a combined model of INT and EXT trajectories between age 7 and 15 was tested, in which the EXT latent categorical class variable (C_{E_i}) was regressed on the INT latent categorical class variable (C_{I_i}) (see Figure 1). This multinomial regression analysis provided an omnibus test of the null hypothesis that INT and EXT classes are unrelated. Note that the direction of this regression is arbitrary, and has no effect on the interpretation of the results. Reversing the direction of the regression to EXT on INT would result in exactly the same model fit and parameter estimates. INT and EXT class variables, denoted C_{I_i} and C_{E_i} , were also regressed on maternal social class, paternal social class, maternal educational level, maternal age at delivery, maternal smoking during pregnancy, birth weight of the child, and sex of the child to test whether these variables predict trajectories.

[Figure 1 about here]

Models were fitted with M-plus 6.12 (Muthén & Muthén, 2007) using robust full information maximum likelihood. If initial settings did not result in replicated minima, the number of starts was increased from 500 to 2000 and the number of final optimization from 50 to 200. If the best likelihood was not replicated with 2000 starts, the model was considered to have failed. The choice of the best fitting model was based on the sample size adjusted BIC. In case of small differences in fit, the preciseness of individual assignment to a specific trajectory and the interpretability of the model were also considered. The certainty of class assignment of the individuals is captured by the entropy

index presented in the results section. A higher entropy implies a higher degree of certainty concerning the assignments.

Previous analyses have showed that missingness in the ALSPAC data is not random, but that it only marginally affects parameter estimates in statistical analyses (Wolke, Waylen, Samara, Steer, Goodman, Ford, & Lamberts, 2009). As a form of attrition analysis we regressed the number of missing DAWBA assessments per individual on the covariates. This attrition analysis showed that sex, smoking during pregnancy, maternal and paternal social class, maternal highest education and maternal age at delivery significantly predicted missingness ($R = 0.274$, $F = 124.6$, $p < 0.0001$). As these variables were included in the model, our model is robust for missingness conditional on these variables (i.e., data missing at random (Little & Rubin 2014)). We reran the models on listwise complete data to evaluate the models based on individuals with complete data, the same conclusions regarding the class selection would be drawn.

Results

Descriptives. Table 1 provides an overview of the prevalences of the observed DAWBA band scores and the polychoric correlations between INT and EXT at age 7, 10, 13 and 15. Polychoric correlations quantify the association between ordinal variables (Ekstrom, 2011). The estimated prevalence of externalizing disorders in our sample was around 5% between ages 7 and 15. The prevalence of internalizing disorders was around 4% in childhood, and grew to 5% at age 15. As expected, male average EXT scores were greater than female average EXT scores at all ages, and female average INT scores were greater than male average INT at ages 13 and 15. Correlations between EXT and INT were around .20. Longitudinal correlations for INT between age 7 and 15 were .15 to .48, whereas correlations for EXT between the ages of 7 and 15 were higher (.35 to .61).

[Table 1]

Internalizing and externalizing trajectories. First, GMMs were fitted for INT and EXT separately. The single class model with a random intercept, slope and quadratic term showed a worse fit than models including a latent class variable, which indicates the existence of subgroups with different trajectories. Models with 2-6 classes were tested with 1) a *fixed* intercept, slope, and quadratic term, and 2) a *random* intercept, and a *fixed* slope and quadratic term. Table 2 provides the model fit statistics and entropy. We retained the quadratic term as models with the quadratic term generally outperformed models without a quadratic term (results available on request from the first author). INT data were best described by a model with 5 classes with a fixed intercept, slope and quadratic term. For EXT, the best fitting model is the 3 class random intercept model. However, this model has a very low entropy compared to the fixed intercept models. Among the fixed intercept models, the best fitting model is the 6 class model, but the 5 class model has a substantially better entropy and only a slightly worse fit. Visual inspection of the trajectories showed that the 6 class model adds a third unaffected class to the very low and low classes, which starts out low and progresses to very low EXT scores. As this extra class is not very informative, the 5 class fixed intercept model was preferred. Results of the analyses of the listwise complete data were similar and also resulted in the selection of a 5 class model for both INT and EXT. Models fitted on listwise complete data had a higher entropy reflecting that individuals with complete data available are easier to categorize.

[Table 2]

Combined internalizing/externalizing model. In the combined model, the association between INT and EXT was analyzed using the multinomial logistic regression of the 5 class EXT trajectories on the 5 class INT trajectories (Figure 1). The model including the multinomial regression parameters fitted the data better than a model that dropped these parameters (Likelihood ratio:

477.894, $df=16$, $p < 0.0001$). We first describe the INT and EXT trajectories and then discuss the association between the EXT and INT trajectory class variables.

[Figure 2]

For INT, there were two classes with low scores, called the very-low-INT class (22.7% of the sample based on most likely class membership) and the low-INT class (41.8%) (Figure 2A). A third class contained individuals with decreasing INT scores (5.1%). The remaining two classes contained individuals with increasing scores. The increasing-INT class (17.8%) showed a steady raise in score from childhood on, while in the adolescent-risk-INT class (12.6%) the scores are low until age 13 but sharply increase at age 15 years.

Four of the five EXT trajectories showed similar patterns as the INT trajectories (Figure 2B), i.e., the very low-EXT class (28%), the low-EXT class (54%), the decreasing-EXT class (7%) and the increasing-EXT class (8.3%). The final high-EXT class (2.4%) was different, as it contained individuals with persisting high scores from childhood to adolescence.

Sex was a significant predictor of INT and EXT class membership, with girls being significantly more likely than boys to be a member of the decreasing (OR= 2.011, $p < 0.001$), increasing (OR= 7.800, $p < 0.001$), or adolescent increasing INT classes (OR=3.128, $p < 0.001$). In addition, girls were significantly less likely to be a member of the high (OR= 0.074, $p < 0.001$), increasing (OR=0.475, $p < 0.001$), and decreasing (OR= 0.178, $p < 0.001$) EXT class. Maternal smoking during pregnancy was a significant risk factor for being a member of the decreasing INT class (OR= 1.991, $p < 0.001$) and for the high (OR= 2.237, $p < 0.001$), increasing (OR= 2.053, $p < 0.001$), and decreasing (OR= 2.765, $p < 0.001$) EXT classes. With respect to EXT, higher social class of the father was associated with a lower probability of belonging to the high (OR= 0.787, $p < 0.001$) or increasing EXT class (OR= 0.818,

$p < 0.001$), and higher maternal education reduced the probability of membership of the increasing EXT class ($OR = 0.790$, $p < 0.001$).

Figure 3 displays the conditional probabilities of belonging to the EXT (INT) classes given membership of a given INT (EXT) class. These conditional probabilities showed that similar internalizing and externalizing classes were associated. Focusing on the “affected” trajectories revealed that individuals in the decreasing INT class had a high probability of belonging to the decreasing EXT class (38%) and children in the increasing INT class had a substantial probability (22%) of being member of the increasing EXT class (Figure 3A). Vice versa, 27% of the children in the decreasing EXT class belonged to the decreasing INT class, and children in the increasing EXT class had a substantial chance (46%) of belonging to the increasing INT class (Figure 3B). It further becomes apparent that the high EXT class was particularly associated with the decreasing INT class and less with the increasing INT class, whereas the adolescent onset INT class was independent from EXT trajectories.

[Figure 3]

Discussion

Based on analyses of DSM-IV based internalizing and externalizing problem scores obtained in longitudinal studies from childhood into adolescence, we conclude that developmental trajectories for INT and EXT are largely similar, and that the INT and EXT trajectories are associated. Notable differences in trajectories are that for INT there is a trajectory that is characterized by increased scores from adolescence onwards, while for EXT there is a trajectory that is characterized by stable high scores. The adolescent onset INT group showed no association with affected EXT classes. The high EXT group was most associated with the decreasing INT group signifying that in some of the

individuals that have both externalizing and internalizing symptoms during childhood, the externalizing symptoms will persist, while the internalizing symptoms attenuate.

Our findings for the INT and EXT trajectories are largely in line with *a priori* expectations. The most apparent discrepancies are the absence of a class characterized by stable high INT and the presence of an increasing EXT class. The absence of a stable high INT group may be due to the relatively low prevalence of these disorders, especially during childhood. Growth mixture models require very large samples to reliably identify classes that consist of a small proportion of the sample. Our findings therefore do not rule out that there is a small group of children with persisting symptoms, as suggested by other studies (Hofstra et al., 2000; Hofstra et al., 2002). Previous results regarding an increasing trajectory for externalizing symptoms were mixed. Van Lier et al. (2007) and Larsson et al. (2011) identified a class with increasing symptoms for conduct disorder, and for the inattentive subtype of ADHD, respectively. The increasing externalizing class in our study probably comprises these groups.

The only two other studies (Brezo et al., 2008; Fanti & Henrich, 2010) that looked at the combination of internalizing and externalizing trajectories up to age 12 also showed that increasing and decreasing internalizing and externalizing trajectories are mutually dependent. It will be interesting to see whether future studies modeling the trajectories into adolescence will replicate our finding that internalizing disorders with onset in adolescence are independent of externalizing disorders and that persisting high EXT is mainly associated with decreasing INT. Our results indicate that the previously reported longitudinal association between internalizing and externalizing disorders starts in childhood. The recent finding in ALSPAC that adolescent depression is predicted by conduct problems in childhood (Stringaris, Lewis & Maughan 2014) might be attributable to persistent childhood internalizing symptoms.

As mixture modeling is an exploratory technique, the results presented here require replication (Lubke, 2012). A related issue is the relatively low entropy, or certainty of class assignments in the models, suggesting that class assignment based on the model is imprecise. We note that entropy was higher in the analyses carried out with listwise complete data: 0.76 for the 5 class externalizing model and 0.572 for the 5 class internalizing model. This reflects the fact that class assignment is substantially more accurate in subjects with data available at each time point, However, while class assignment is more precise for individuals with complete data, including subjects with missing data has the advantage of rendering the results robust to missingness associated with a high score at an earlier measurement occasion, and with an included risk factor. Although we fit a model robust for missingness associated with the included covariates, we acknowledge dropout associated with other covariates, not included in our models, may have affected the results in an unknown way. We note that the associations between sex, prenatal risk factors and trajectory membership probability were all in the expected directions. Boys were found to be more at risk for externalizing problems and girls more for internalizing problems and adverse prenatal risk factors were associated with externalizing, and to a lesser extent, with internalizing problems (see e.g., (Ormel, Raven, van Oort, Hartman, Reijneveld, Veenstra, Vollebergh, Buitelaar, Verhulst, & Oldehinkel, 2014). In sum, the agreement between the estimated trajectories and expectations based on previous work increases the confidence in these developmental trajectories.

One of the other potential limitations, besides the attrition rate, could be the use of two broad internalizing and externalizing problem scores. Studies investigating specific internalizing or externalizing symptom domains (Barker, Séguin, White, Bates, Lacourse, Carbonneau, & Tremblay, 2007; Broeren et al., 2013; Larsson et al., 2011; Van Lier et al., 2007; Broeren, Muris, Diamantopoulou, & Baker, 2013) have detected differences in trajectories between the separate disorders. Given the low prevalence rates of the individual disorders, such analyses were not feasible here. Moreover, previous studies have also shown that the analyzed disorders load on common

factors interpretable as our INT and EXT (see e.g., Angold et al., 1999; Beauchaine & McNulty, 2013; Cosgrove et al., 2011). This indicates that studies focusing on measures of a general tendency to display internalizing or externalizing disorders can also provide important information. As sample size precluded separate analyses in males and females, gender was included as a covariate predicting class membership. We acknowledge that the slight variation in constituent disorders and raters over time is a limitation. Internalizing disorders were measured by self rating at age 15, and by maternal rating at ages 7 to 13. The design of the ALSPAC study is such that the set of internalizing disorders at age 7 to 13 includes separation anxiety, while the measure at age 15 includes agoraphobia and panic disorder. This is based on the expectation that agoraphobia in early childhood and separation anxiety in late adolescence are unlikely to be present. The change in disorders measured across the time span of this study can be viewed as an attempt to measure heterotypic continuity of the same underlying internalizing construct, a technique often applied when considering longitudinal development (Petersen, Bates, Dodge, Lansford, & Pettit, 2014). Finally, the correlations between internalizing and externalizing disorders (around 0.20) were lower than previously reported by Cosgrove et al. (2011) (0.20 to 0.30). However, these differences are relatively small, and might be due to differences in the instrument used.

Other questionnaire disorders, such as the child behavior checklist (CBCL) and the youth self report (YSR) (Achenbach & Rescola, 2001) are often used to measure internalizing and externalizing psychopathology in large population samples. Growth curve analyses based on the CBCL/YSR internalizing symptoms yielded evidence for heterogeneous development of internalizing problems, but this variation could not be easily disaggregated into different latent growth classes, i.e., a model including a random intercept and a slope provided the best fit (Lubke et al. 2015). This difference in outcome is likely to be due the fact that the CBCL is a continuous measure capturing individual differences in the general population, and as such provides more information to estimate random

growth parameters, whereas the measure used in this study is ordinal and developed for clinical practice.

The strengths of this study were the use of a large population based sample with repeated measures from childhood into adolescence and, importantly, a DSM-IV based psychiatric interview instrument. This enables the translation of these findings to clinical relevance. The results suggest that if an adolescent presents with internalizing symptoms and has no history of previous internalizing symptoms, a brief screening for externalizing disorders will suffice. However, when confronted with childhood or adolescent externalizing problems or with internalizing problems that were already apparent in childhood, co-morbid symptoms should also be assessed at the start of the treatment. Risk and protective factors can also be taken into account. If the mother smoked during pregnancy, the chance to have a trajectory of increasing or persisting externalizing symptoms is higher, whereas internalizing symptoms may decrease. A protective factor for unfavorable EXT trajectories is higher social class. Future studies should address the specific treatment needs of children with co-occurring internalizing and externalizing disorders, especially since the co-occurrence is related to negative outcomes (Fanti & Henrich, 2010). An interesting question is whether successful treatment of an externalizing disorder also leads to a remission of the internalizing disorder, or vice versa, or whether treatment of both disorders is necessary. There is some evidence that internalizing symptoms respond to the treatment of externalizing symptoms (Chase & Eyberg, 2008), and vice versa (Kendall, Brady & Verduijn, 2001). Moreover, it is important to identify the factors associated with the combination of the trajectories of decreasing internalizing and externalizing symptoms versus the combination of persisting externalizing symptoms and decreasing internalizing symptoms.

The current study does not address the etiology of comorbidity between internalizing and externalizing disorders. Different hypotheses currently exist about the causes of comorbidity. It has

been suggested that depressive symptoms in ADHD are due to demoralization (Brown, Borden, Clingerman, & Jenkins, 1988), but in line with our finding that a combination of trajectories of externalizing symptoms in childhood and later increasing internalizing symptoms did not exist, Biederman, Mick, Faraone (1998) concluded that this does not explain all comorbidity. The opposite, i.e. internalizing symptoms causing externalizing symptoms, has also been hypothesized. Granic (2014), for example, proposes three mechanisms explaining how anxiety can cause aggression and suggests how future research could investigate whether these mechanisms play a role. Another explanation for comorbidity is that multiple disorders are caused by the same underlying mechanism, which agrees with the observed co-occurrence of similar trajectories. Cross-sectional twin studies have indicated that co-morbidity between internalizing and externalizing disorders is partly explained by shared genetic risk factors (e.g. Cosgrove et al. 2011). It has previously been shown that ADHD trajectories are influenced by genetic factors (Larsson et al., 2011). This could also be the case for co-occurring trajectories, which would be interesting for gene-finding studies. Including a genetic variant-by-course (i.e., decreasing or stable high) interaction term enables the identification of variants associated with a favorable or unfavorable outcome and reveal hints about biological differences in etiology between developmental courses.

The mechanisms underlying the association between the risk factors, such as maternal smoking during pregnancy, paternal social class and education, and the developmental trajectories are also a subject for further investigation. Many of the associated risk factors are influenced by genetic factors. The question whether the associations between exposures and adverse trajectories are explained by common genetic factors, common environmental exposures or by direct causation can be addressed by genetically informative designs, such as studies of developmentally concordant and discordant monozygotic and dizygotic twins (van Dongen, Slagboom, Draisma, Martin & Boomsma, 2012; McGue, Olsen & Christensen, 2010) or children-of-twins Designs (McAdams, Neiderhiser, Rijdsdijk, Narusyte, Lichtenstein & Eley 2014). The advent of high throughput genome

wide (epi)genetic measurement will further allow for a more in-depth study of the origins of biological differences underlying developmental differences. Specifically, subtle developmental differences beyond lifetime diagnoses are interesting targets for studies relating epi-genetic variation to variations in psychopathology.

To summarize, we showed that both internalizing and externalizing disorders can have a favorable or unfavorable course in time from childhood into adolescence and that trajectories are associated with each other. Future research should focus on unraveling the etiology of the co-occurrence, and focus on the development of treatment designs for the most seriously affected children.

Reference List

- Achenbach, T., & Rescorla, L. A. Manual for the ASEBA School-Age Forms & Profiles. 2001, University of Vermont. *Research Center for Children, Youth, & Families, Burlington.*
- Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40*, 57-87.
- Barker, E. D., Séguin, J. R., White, H. R., Bates, M. E., Lacourse, E., Carbonneau, R. Tremblay, R.E. (2007). Developmental trajectories of male physical violence and theft: relations to neurocognitive performance. *Archives of General Psychiatry, 64*, 592-599.
- Beauchaine, T. P. & McNulty, T. (2013). Comorbidities and continuities as ontogenic processes: Toward a developmental spectrum model of externalizing psychopathology. *Development and psychopathology, 25*, 1505-1528.
- Biederman, J., Mick, E., & Faraone, S. V. (1998). Depression in attention deficit hyperactivity disorder (ADHD) children: "True" depression or demoralization? *Journal of Affective Disorders, 47*, 113-122.
- Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S. & Smith, G. D. (2012). Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International journal of epidemiology, dys064*.

- Brezo, J., Barker, E. D., Paris, J., Hébert, M., Vitaro, F., Tremblay, R. E., Turecki, G. (2008). Childhood trajectories of anxiousness and disruptiveness as predictors of suicide attempts. *Archives of pediatrics & adolescent medicine*, *162*, 1015-1021.
- Broeren, S., Muris, P., Diamantopoulou, S., & Baker, J. R. (2013). The course of childhood anxiety symptoms: developmental trajectories and child-related factors in normal children. *Journal of Abnormal Child Psychology*, *41*, 81-95.
- Brown, R. T., Borden, K. A., Clingerman, S. R., & Jenkins, P. (1988). Depression in attention deficit-disordered and normal children and their parents. *Child psychiatry and human development*, *18*, 119-132.
- Chase, R. M., & Eyberg, S. M. (2008). Clinical presentation and treatment outcome for children with comorbid externalizing and internalizing symptoms. *Journal of Anxiety Disorders*, *22*, 273-282.
- Copeland, W. E., Shanahan, L., Costello, E. J., & Angold, A. (2009). Childhood and adolescent psychiatric disorders as predictors of young adult disorders. *Archives of General Psychiatry*, *66*, 764-772.
- Cosgrove, V. E., Rhee, S. H., Gelhorn, H. L., Boeldt, D., Corley, R. C., Ehringer, M. A., Young, S.E., Hewitt, J.K.(2011). Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. *Journal of abnormal child psychology*, *39*, 109-123.
- Costello, E. J., Copeland, W., & Angold, A. (2011). Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? *Journal of Child Psychology and Psychiatry*, *52*, 1015-1025.
- Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *Jama*, *290*, 2023-2029.

- Dekker, M. C., Ferdinand, R. F., Van Lang, N. D., Bongers, I. L., van der Ende, J., & Verhulst, F. C. (2007). Developmental trajectories of depressive symptoms from early childhood to late adolescence: gender differences and adult outcome. *Journal of Child Psychology and Psychiatry, 48*, 657-666.
- Dolan, C. V., Geels, L., Vink, J. M., van Beijsterveldt, C. E. M., Neale, M. C., Bartels, M., & Boomsma, D. I. (2015). Testing Causal Effects of Maternal Smoking During Pregnancy on Offspring's Externalizing and Internalizing Behavior. *Behavior genetics, 1-11*.
- Ekstrom, J. (2011). *A generalized definition of the polychoric correlation coefficient*. Department of Statistics, UCLA.
- Fanti, K. A. & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: findings from the National Institute of Child Health and Human Development Study of Early Child Care. *Developmental psychology, 46*, 1159-1175.
- Golding, J., Pembrey, M., & Jones, R. (2001). ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatric and Perinatal Epidemiology, 15*, 74-87.
- Goodman, A., Heiervang, E., Collishaw, S., & Goodman, R. (2011). The DAWBA bands as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. *Social psychiatry and psychiatric epidemiology, 46*, 521-532.
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology. *Journal of Child Psychology and Psychiatry, 41*, 645-655.

- Granic, I. (2014). The role of anxiety in the development, maintenance, and treatment of childhood aggression. *Development and psychopathology, 26*, 1515-1530.
- Groen-Blokhuis, M. M., Middeldorp, C. M., van Beijsterveldt, C. E., & Boomsma, D. I. (2011). Evidence for a causal association of low birth weight and attention problems. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*, 1247-1254.
- Hack, M., Youngstrom, E. A., Cartar, L., Schluchter, M., Taylor, H. G., Flannery, D., ... & Borawski, E. (2004). Behavioral outcomes and evidence of psychopathology among very low birth weight infants at age 20 years. *Pediatrics, 114*(4), 932-940.
- Haltigan, J. D., Roisman, G. I., Susman, E. J., Barnett-Walker, K., & Monahan, K. C. (2011). Elevated trajectories of externalizing problems are associated with lower awakening cortisol levels in midadolescence. *Developmental psychology, 47*, 472-478.
- Hofstra, M. B., Van der Ende, J., & Verhulst, F. C. (2000). Continuity and change of psychopathology from childhood into adulthood: a 14-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 850-858.
- Hofstra, M. B., Van der Ende, J., & Verhulst, F. C. (2002). Child and Adolescent Problems Predict DSM-IV Disorders in Adulthood: A 14-Year Follow-up of a Dutch Epidemiological Sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*, 182-189.
- Jung, T. & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass, 2*, 302-317.
- Kendall, P. C., Brady, E. U., & Verduin, T. L. (2001). Comorbidity in childhood anxiety disorders and treatment outcome. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*, 787-794.

- Larsson, H., Dilshad, R., Lichtenstein, P., & Barker, E. D. (2011). Developmental trajectories of DSM IV symptoms of attention deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology. *Journal of Child Psychology and Psychiatry*, *52*, 954-963.
- Letcher, P., Smart, D., Sanson, A., & Toumbourou, J. W. (2009). Psychosocial Precursors and Correlates of Differing Internalizing Trajectories from 3 to 15 Years. *Social Development*, *18*, 618-646.
- Linnet, K. M., Dalsgaard, S., Obel, C., Wisborg, K., Henriksen, T. B., Rodriguez, A., Kotimaa, A., Moilanen, I., Hove Thomsen, P., Olsen, J. & Jarvelin, M. R. (2003). Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: review of the current evidence. *American Journal of Psychiatry*, *160*, 1028-1040.
- Little, R. J., & Rubin, D. B. (2014). *Statistical analysis with missing data*. John Wiley & Sons.
- Lubke, G. (2012). Old Issues in a New Jacket: Power and Validation in the Context of Mixture Modeling. *Measurement: Interdisciplinary Research and Perspectives*, *10*, 212-216.
- Lubke, G. H., Miller, P. J., Verhulst, B., Bartels, M., van Beijsterveldt, T., Willemsen, G., Boomsma, D.I., & Middeldorp, C. M. (2015). A powerful phenotype for gene-finding studies derived from trajectory analyses of symptoms of anxiety and depression between age seven and 18. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*.
- McAdams T.A., Neiderhiser J.M., Rijdsdijk F.V., Narusyte J., Lichtenstein P., Eley T.C. (2014) Accounting for genetic and environmental confounds in associations between parent and child characteristics: a systematic review of children-of-twins studies. *Psychological Bulletin*. *140*, 1138-1173

- McGue, M., Osler, M., & Christensen, K. (2010). Causal inference and observational research the utility of twins. *Perspectives on Psychological Science, 5*, 546-556.
- Muthen, B. & Muthen, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and experimental research, 24*, 882-891.
- Muthén, L. K. & Muthén, B. O. (2007). *Mplus User's Guide. Fifth Edition*. Los Angeles, CA: Muthén & Muthén.
- Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological methods, 4*, 139-157.
- Newman, D., Moffitt, T., Caspi, A., & Silva, P. (1998). Comorbid mental disorders: Implications for treatment and sample selection. *Journal of Abnormal Psychology, 107*, 305–311
- Ormel, J., Raven, D., van Oort, F., Hartman, C. A., Reijneveld, S. A., Veenstra, R. Vollebergh, W.A., Buitelaar, J., Verhulst, F.C., Oldehinkel, A.J. (2014). Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. *27*, 345-360.
- Petersen, I. T., Bates, J. E., Dodge, K. A., Lansford, J. E., & Pettit, G. S. (2015). Describing and predicting developmental profiles of externalizing problems from childhood to adulthood. *Development and psychopathology, 27*, 791-818.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

- Stringaris, A., Lewis, G., & Maughan, B. (2014). Developmental pathways from childhood conduct problems to early adult depression: findings from the ALSPAC cohort. *The British Journal of Psychiatry*, 205, 17-23.
- Toumbourou, J. W., Williams, I., Letcher, P., Sanson, A., & Smart, D. (2011). Developmental trajectories of internalising behaviour in the prediction of adolescent depressive symptoms. *Australian journal of psychology*, 63, 214-223.
- van Dongen, J., Slagboom, P. E., Draisma, H. H., Martin, N. G., & Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. *Nature Reviews Genetics*, 13, 640-653.
- Van Lier, P. A., Der Ende, J. v., Koot, H. M., & Verhulst, F. C. (2007). Which better predicts conduct problems? The relationship of trajectories of conduct problems with ODD and ADHD symptoms from childhood into adolescence. *Journal of Child Psychology and Psychiatry*, 48, 601-608.
- Weissman, M. M., Warner, V., Wickramaratne, P. J., & Kandel, D. B. (1999). Maternal smoking during pregnancy and psychopathology in offspring followed to adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38, 892-899.
- Wiggins, J. L., Mitchell, C., Hyde, L. W., & Monk, C. S. (2015). Identifying early pathways of risk and resilience: The codevelopment of internalizing and externalizing symptoms and the role of harsh parenting. *Development and Psychopathology*, 1-18.
- Wolke, D., Waylen, A., Samara, M., Steer, C., Goodman, R., Ford, T. et al. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *The British Journal of Psychiatry*, 195, 249-256.

Table 1: For each age, the number of individuals for each DAWBA band score for externalizing (EXT) and internalizing (INT) disorders, prevalence's for EXT and INT and total N (upper part), and the polychoric correlations (all significant, $p < 0.05$ corrected for multiple testing.) between EXT and INT (lower part).

Descriptives

	DAWBA band score					Prevalence	N total
	1	2	3	4	5		
EXT 7	1509	3948	612	160	96	0.058	6325
EXT 10	1668	3388	520	124	90	0.054	5790
EXT 13	1848	2682	608	115	79	0.055	5332
EXT 15	1714	1848	335	88	65	0.05	4050
INT 7	2749	3063	412	76	45	0.037	6345
INT 10	2092	3101	470	89	58	0.045	5810
INT 13	2181	2686	382	68	33	0.038	5350
INT 15	1506	1885	521	106	15	0.051	4033

*Correlations (below
diagonal /SE (above
diagonal))*

	EXT 7	EXT 10	EXT 13	EXT 15	INT 7	INT 10	INT 13	INT 15
EXT 7	1	0.009	0.0011	0.015	0.012	0.014	0.0016	0.019
EXT 10	0.63	1	0.01	0.014	0.015	0.013	0.015	0.019
EXT 13	0.497	0.582	1	0.011	0.016	0.015	0.013	0.018
EXT 15	0.387	0.451	0.569	1	0.019	0.018	0.018	0.016
INT 7	0.23	0.161	0.102	0.083	1	0.011	0.013	0.019
INT 10	0.188	0.257	0.151	0.133	0.485	1	0.011	0.018
INT 13	0.162	0.177	0.211	0.173	0.406	0.483	1	0.018
INT 15	0.081	0.08	0.074	0.11	0.142	0.19	0.225	1

N classes	Fixed ISQ				Random I, Fixed SQ,				Random ISQ	
	Entropy	AIC	Adj-BIC	LMR	Entropy	AIC	Adj-BIC	LMR	AIC	Adj-BIC
1									41748.603	41870.892
2	0.469	42326.785	42389.782	P < 0.001	0.323	41946.080	42016.489	P < 0.001	-	
3	0.432	41959.951	42063.711	P < 0.001	0.455	41674.774	41789.651	P < 0.001	-	
4	0.513	41650.452	41794.975	P < 0.001	0.392	41536.956	41696.303	P < 0.001	-	
5	0.488	41536.394	41721.680	P < 0.001		*			-	
6	0.518	41505.755	41731.804	P = 0.021		*			-	

Table 2: Fit indices for the internalizing (INT) and externalizing (EXT) growth mixture models

containing 1 to 6 classes. On the left, the models with fixed effects for intercept (I), linear (S) and quadratic (Q) slopes i.e., with the variances of the intercept and slopes fixed to zero in each class. In the middle models with a random I (i.e., with the variance of the intercept estimated in each class) and fixed effects for S and Q. And on the upper right the reference model with random I, S and Q. The LMR column provides the p-value for the Lo-Mendell-Rubin test, which tests the appropriately adjusted likelihood ratio between the model under consideration and the model with 1 class less.

* best log-likelihood not replicated at 2000 starts and 200 final iterations.

Models EXT

N classes	Fixed ISQ				Random I, Fixed SQ,				Random ISQ	
	Entropy	AIC	Adj-BIC	LMR	Entropy	AIC	Adj-BIC	LMR	AIC	Adj-BIC
1									41273.385	41395.628
2	0.544	42770.286	42833.259	P < 0.001	0.383	41463.420	41533.803	P < 0.001	-	
3	0.647	41630.458	41734.179	P < 0.001	0.471	41235.535	41350.369	P < 0.006	-	
4	0.647	41355.296	41499.765	P < 0.001		*			-	
5	0.656	41208.013	41393.230	P < 0.002		*			-	
6	0.569	41136.539	41362.503	P < 0.001		*			-	

Figure legends:

Figure 1.: Path model of the final growth mixture model. C_i and C_e are the class variables indicating the distinct growth trajectories between the ages of 7 and 15 years for the internalizing and externalizing problem scores, respectively. Class membership of internalizing problems is modeled to predict class membership of externalizing problems. For each growth trajectory class the means of the intercept (i), linear slope (s) and quadratic slope (q) are estimated. The factor loadings of the intercept (i) are fixed to 1. The factor loadings of the linear slope (s) are fixed to 1, 2 and 2.66, and of the quadratic slope (q) 1, 4 and 7.07, respectively. These constants are proportional to the differences between the measurement occasions expressed in years, or years squared.

Subscript i : internalizing disorders, subscript e : externalizing disorders

Figure 2. A) The five internalizing trajectories. B) The five externalizing trajectories. The y axis indicates the expected DAWBA band score for a given class at a given age.

Figure 3: A) Probabilities for EXT class membership and conditional probabilities for EXT class membership given INT class membership. The bars at the left, above "all" indicate the probabilities to belong to the five EXT classes as estimated in the whole sample. Next, the bars above "very low INT" indicate the probabilities to belong to the five EXT classes as estimated for the children that were assigned to the very low INT class. The same applies to the bars in the categories Low INT, Increasing INT, Decreasing INT and Ado-increasing INT. Clearly, the children in the very low and low EXT classes are overrepresented in the low INT classes and in the class with adolescent increasing INT, whereas the children with increasing or high EXT are overrepresented in the increasing and decreasing internalizing classes.

B) Probabilities for INT class membership and conditional probabilities for INT class membership given EXT class membership. Clearly, the children in the very low and low INT classes are over

represented in the children in the low EXT classes, whereas the children with increasing or decreasing INT are overrepresented in the increasing, decreasing and high EXT classes.”

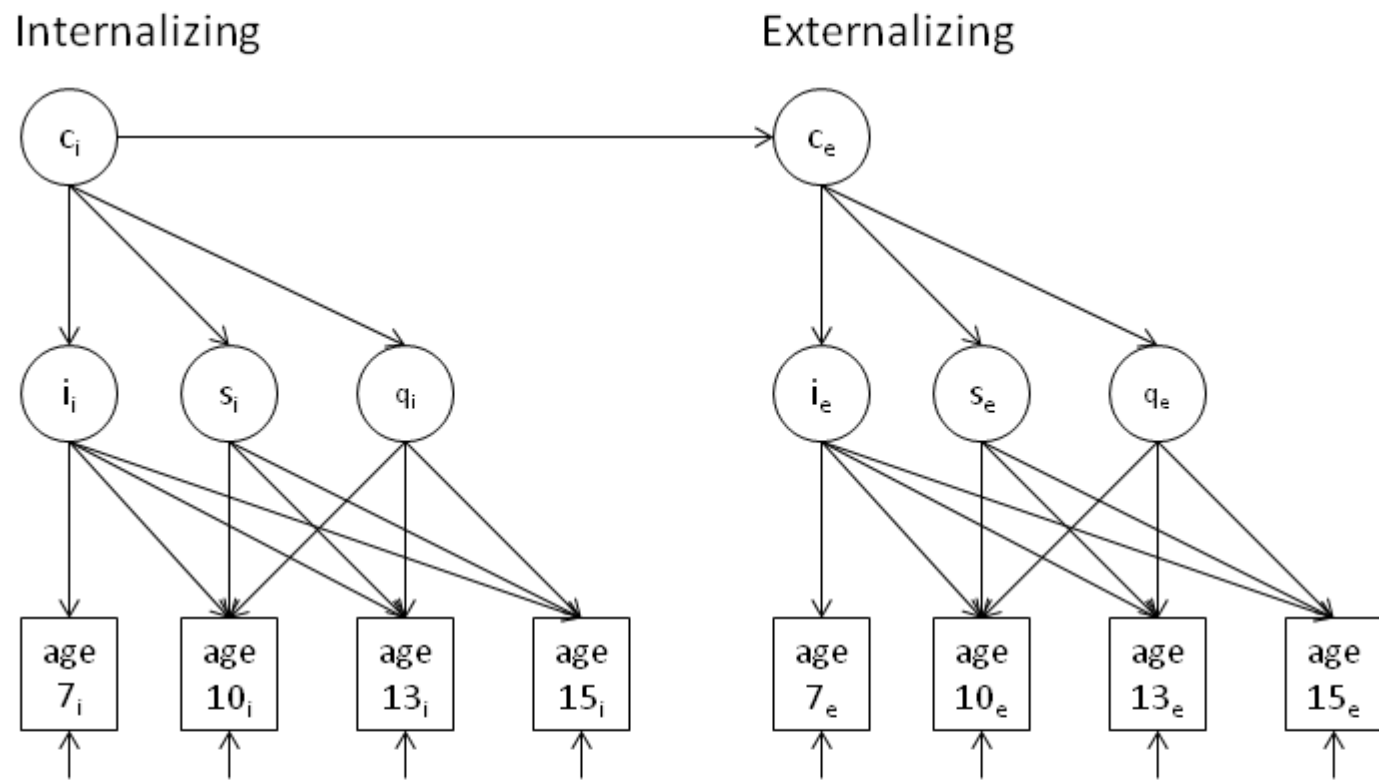
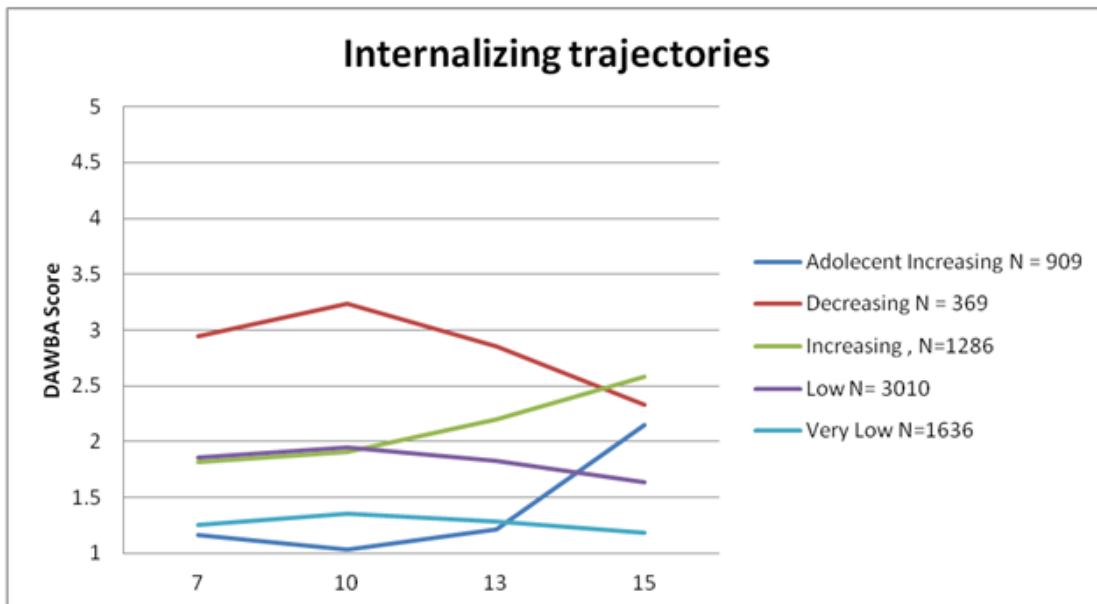


Figure 1

A)



B)

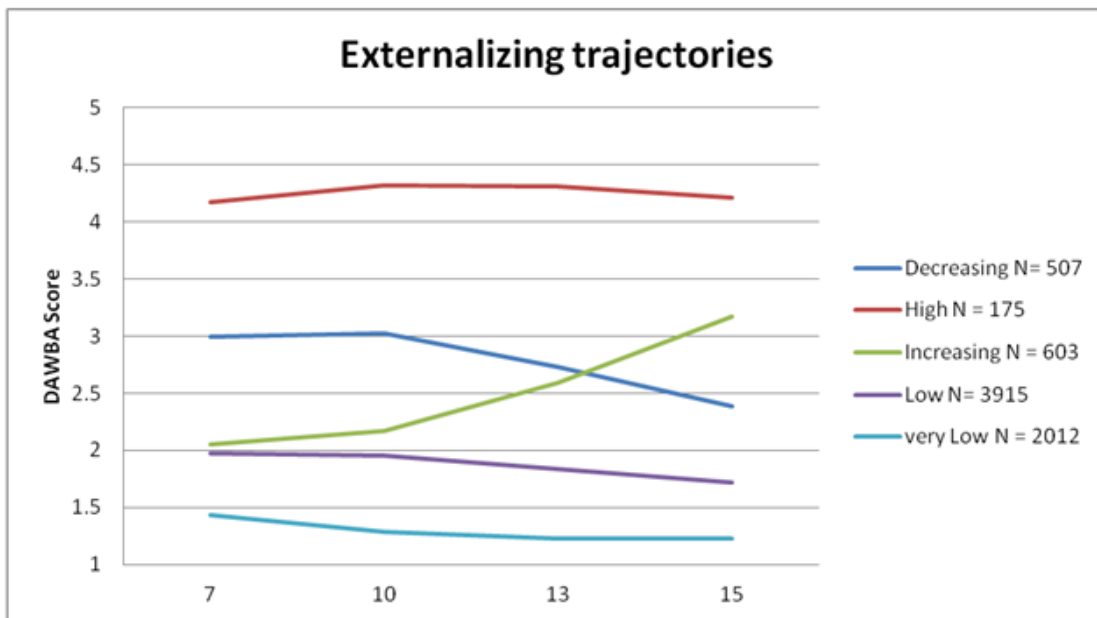
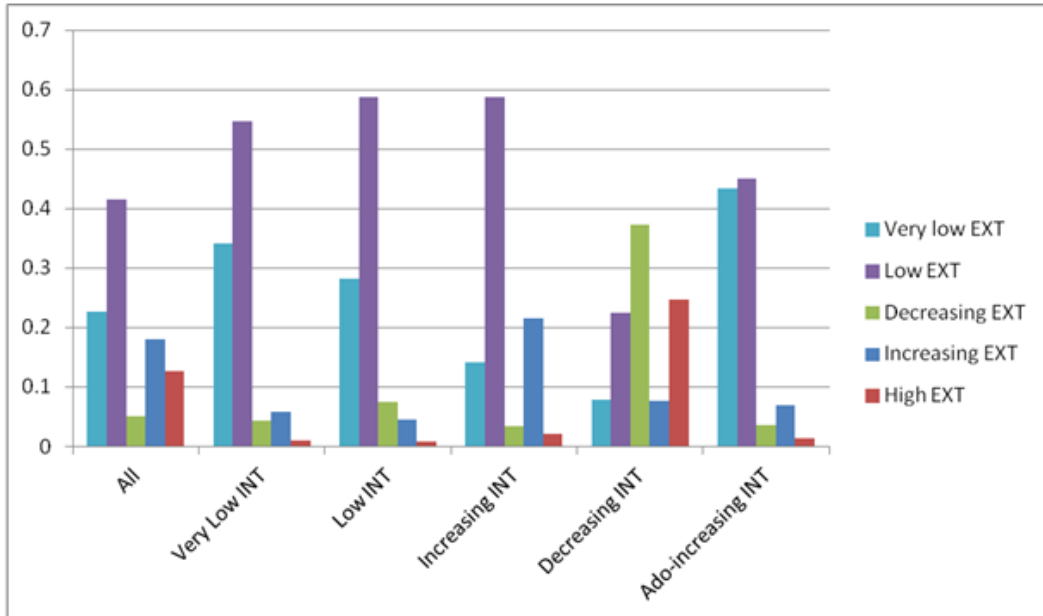


Figure 2

A)



B)

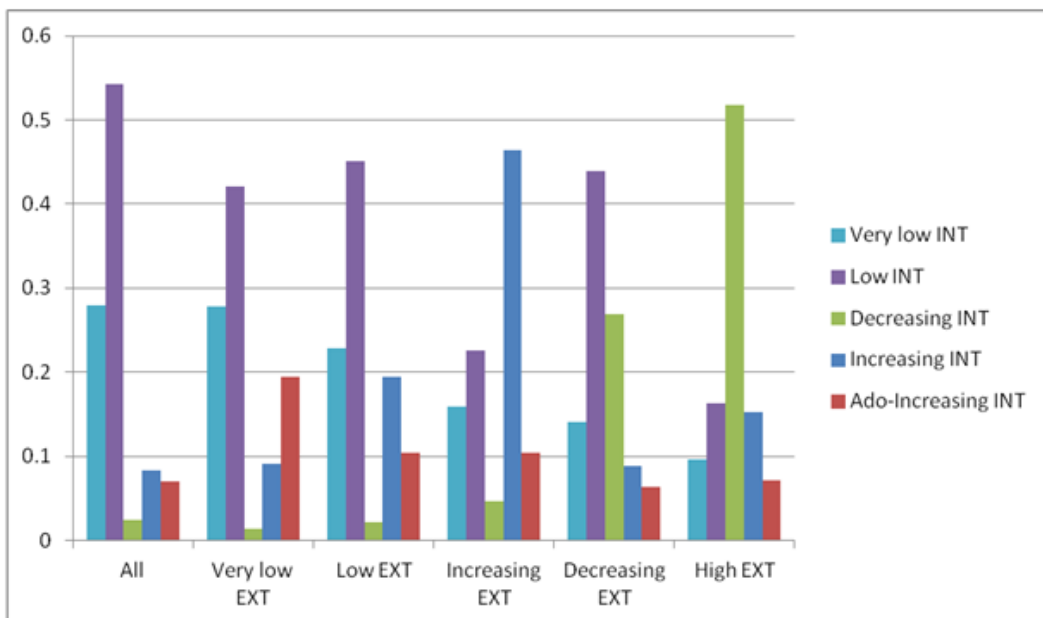


Figure 3