**Online supplement**

**The development of latent dimensions of psychopathology across early childhood:**

**Stability of dimensions and moderators of change**

**Expanded Description of the Laboratory Assessment**

Temperament was assessed in 541 children during a 2-hour laboratory visit which included a structured observation consisting of 12 episodes designed to elicit a range of temperament-related behaviors. Eleven of these episodes were taken from the Laboratory Temperament Assessment Battery ([Lab-TAB; Goldsmith, Reilly, Lemery, Longley, & Prescott, 1995](#_ENREF_31)), and one was adapted from a Lab-TAB episode. The Lab-TAB has shown moderate convergent validity with post-visit observer ratings (*r*s = .21-.76) ([Gagne, Van Hulle, Aksan, Essex, & Goldsmith, 2011](#_ENREF_30)). In a study of the ecological validity of the Lab-TAB, parents observing their children during the activities tended to rate their child’s behavior as “extremely typical” or “highly typical” of their behavior outside the lab ([Lo, Vroman, & Durbin, 2015](#_ENREF_53)). In an independent sample, we found moderate stability of laboratory ratings of temperament traits from ages 3-7 and moderate concurrent and longitudinal associations with home observations of temperament ([Durbin, Hayden, Klein, & Olino, 2007](#_ENREF_20)). Specific laboratory episodes are described in previous publications (e.g., Olino et al., 2010).

Coding procedures followed those reported in a previous study ([Durbin, Klein, Hayden, Buckley, & Moerk, 2005](#_ENREF_21)). Multiple temperament traits, emotional displays, and behaviors were rated including positive affect (PA), fear, sadness, anger, behavioral inhibition (BI), inhibitory control, interest, activity, sociability, impulsivity, initiative, anticipatory PA, warmth, social interest, affiliation, assertiveness, dominance, pushiness, hostility, noncompliance, avoidance, clinginess, and social fear. Coefficient alphas for internal consistency ranged from .50 to .87 (*Mdn* = .70) and ICCs for interrater reliability ranged from .40 to .92 (*Mdn =* .75, *n* = 35).

In a previous report using this sample, Principal Components Analysis (PCA) of Lab-TAB variables was conducted to reduce the number of temperament variables ([Dougherty, et al., 2011](#_ENREF_17)). This analysis yielded five temperament scales: Sociability/Assertiveness (α = .93), Dysphoria (α = .80), Fear/Inhibition (α = .71), Exuberance (α = .88), and Disinhibition (α = .70). Each of these components included multiple theoretically relevant variables with loadings greater than .4. In the few cases where there cross-loadings, the item was added to the scale with the higher loading such that no items were double-counted. Interrater reliability was adequate for all scales, with interrater ICCs as follows: Sociability/Assertiveness (.82), Dysphoria (.88), Fear (.82), Exuberance (.92), and Disinhibition (.83). See Dougherty et al. (2011) for a complete description.

**Measurement Invariance**

The utility of testing measurement invariance is highlighted when investigators are interested in comparing mean levels of constructs across meaningful units and requires demonstrating configural, metric, and, at least partial, scalar invariance to support those comparisons. In the context of the present work, our emphasis is squarely on the consistency in rank-ordering across time. Thus, the burden of these analyses is to support, at least partial, metric invariance, such that the meaning of our constructs is consistent over time.

Within the model specifying longitudinal correlations between all latent factors, we applied constraints on factor loadings across time to evaluate metric invariance. The model was an adequate fit to the data as indicated by the RMSEA = .069, but inadequate according to other indices, χ2 (120) = 473.79, *p* < .001 and CFI = .868. Current conventions suggest that invariance is supported when there are smaller than .015 reduction in RMSEA and/or CFI ([Chen, 2007](#_ENREF_13)). The reduction in RMSEA from the initial model to the metric invariance model was modest, but the reduction in model fit was marked for the CFI. Comparisons of individual factor loadings (Table S1) indicated that there were non-significant differences in factor loadings for all except the factor loading to inattention symptoms from the common factor, χ2 (1) = 1.26, *p* < .001. We re-estimated the model imposing all equality constraints on factor loadings across time, except for the loading to inattention from the common factors. This partial metric invariance model was an improved fit to the data, relative to the full metric invariance model as indicated by the RMSEA = .067, but inadequate according to other indices, χ2 (119) = 408.78, *p* < .001 and CFI = .879. Similarly, relative to the initial longitudinal mode, the partial did not evince reduced model fit according to the RMSEA, but was reduced according to the CFI. Thus, we did not proceed to test additional layers of invariance. We tested longitudinal prediction of age 6 factors from age 3 factors without imposing measurement invariance.

**Complete Results of Moderation Analyses.**

We estimated a series of models to examine whether child sex or temperament, reports of parenting, or parental history of psychopathology moderated homotypic and heterotypic continuity. Consistent with the previous model, we estimated all homotypic and heterotypic continuity paths simultaneously. Likewise, we estimated all interaction effects simultaneously in the same model. We focused on 12 independent predictors (child sex, 5 dimensions of child temperament, 3 forms of parental psychopathology, and 3 dimensions of maternal parenting behaviors). In each of our models, there were 9 interaction terms estimated simultaneously, for a total of 108 interactions across all models (including both the models for continuous and categorical moderators). Of these interactions, only two reached conventional levels of significance. Factors that did not influence either homotypic or heterotypic continuity of psychopathology include child sex, child temperament (exuberance, disinhibition, dysphoria, and sociability), parental history of depression and anxiety, and maternal authoritative, authoritarian, and maternal permissive parenting. Below, we focus on the significant moderators of homotypic and heterotypic continuity.

Homotypic continuity of the externalizing factor was significantly moderated by child temperamental parental history of SUD and fearful inhibition (see Table S2 and Supplementary Figure 1, top and bottom panels, respectively). Stability of externalizing problems was stronger for youth with higher levels of fearfulness (b = .66, SE = .09, *p* < .001) than lower levels of fearfulness (b = .44, SE = .10, *p* < .01). Stability of externalizing problems was also stronger for youth with one or more parents with a history of SUD (b = .87, SE = .17, *p* < .001) than for youth without a parental history of SUD (b = .36, SE = .12, *p* < .01).

Table S1. Comparisons of Unstandardized Factor Loadings Across Ages 3 and 6

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Factor | Symptom Dimension | Age 3 Factor Loading | Age 6 Factor Loading | *t* |
| Common | DEP | .41 | .30 | 1.66 |
|  | GAD | .42 | .40 | .389 |
|  | SEP | .09 | .09 | .00 |
|  | PANIC | -.00 | .02 | -.53 |
|  | INAT | .33 | .20 | 3.20\*\* |
|  | HYP | .16 | .21 | -1.18 |
|  | IMP | .10 | .09 | .02 |
|  | ODD | .22 | .25 | -.77 |
| Internalizing | DEP | .29 | -.13 | 1.89 |
|  | GAD | .15 | -.19 | -.62 |
|  | SEP | .18 | -.13 | .75 |
|  | PANIC | .15 | -.10 | 1.36 |
|  | PHOB | .20 | -.28 | -1.63 |
| Externalizing | INAT | .26 | .22 | 1.01 |
|  | HYP | .35 | .34 | .38 |
|  | IMP | .22 | .19 | .81 |
|  | ODD | .20 | .25 | -1.07 |

Table S2. Moderators of homotypic and heterotypic stability paths for continuous moderators.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Age 6 General factor | Age 6 Internalizing | Age 6 Externalizing |
| **Child Temperament** |  |  |  |
| Disinhibition | .02 (.01) | .00 (.00) | .01 (.01) |
| Age 3 General factor | .68 (.22)\*\*\* | -.11 (.21) | .1 (.07) |
| Age 3 Internalizing | .4 (1.5) | .77 (.55) | -.07 (.17) |
| Age 3 Externalizing | .24 (.29) | .11 (.15) | .48 (.08)\*\*\* |
| Disinhibition X Age 3 General factor | -.06 (.11) | -.02 (.03) | .04 (.02) |
| Disinhibition X Age 3 Internalizing | .31 (.87) | -.02 (.05) | -.04 (.06) |
| Disinhibition X Age 3 Externalizing | -.14 (.16) | -.01 (.03) | .01 (.04) |
| Dysphoria | .01 (.01) | .01 (0) | .00 (.00) |
| Age 3 General factor | .69 (.19)\*\*\* | -.14 (.14) | .13 (.24) |
| Age 3 Internalizing | -.18 (5.82) | .79 (.34)\* | -.06 (.14) |
| Age 3 Externalizing | .13 (.24) | .09 (.1) | .54 (.08)\*\*\* |
| Dysphoria X Age 3 General factor | -.04 (.04) | -.03 (.03) | -.01 (.03) |
| Dysphoria X Age 3 Internalizing | .04 (.39) | .05 (.05) | -.02 (.04) |
| Dysphoria X Age 3 Externalizing | -.11 (.06) | 0 (.02) | 0 (.03) |
| Exuberance | .01 (.01) | .00 (.00) | .00 (.00) |
| Age 3 General factor | .69 (.15)\*\*\* | -.12 (.11) | .15 (.1) |
| Age 3 Internalizing | .04 (1.53) | .75 (.28)\*\* | -.09 (.15) |
| Age 3 Externalizing | .2 (.19) | .09 (.08) | .55 (.08)\*\*\* |
| Exuberance X Age 3 General factor | -.01 (.02) | .01 (.01) | -.01 (.01) |
| Exuberance X Age 3 Internalizing | .02 (.1) | -.06 (.04) | 0 (.02) |
| Exuberance X Age 3 Externalizing | -.01 (.03) | -.01 (.01) | -.03 (.02) |
| Fear/Inhibition (Fear) | 0 (.01) | .00 (.00) | .01 (0) |
| Age 3 General factor | .69 (.18)\*\*\* | -.18 (.12) | .17 (.09) |
| Age 3 Internalizing | -.17 (2.3) | .90 (.33)\*\* | -.15 (.16) |
| Age 3 Externalizing | .18 (.22) | .12 (.09) | .55 (.08)\*\*\* |
| Fear X Age 3 General factor | .03 (.03) | -.04 (.02)\* | .03 (.02) |
| Fear X Age 3 Internalizing | -.03 (.25) | .08 (.06) | .04 (.04) |
| Fear X Age 3 Externalizing | -.03 (.06) | .04 (.02) | .05 (.03) |
| Sociability/Assertiveness (Soc) | .00 (.00) | .00 (.00) | .00 (.00) |
| Age 3 General factor | .70 (.16)\*\*\* | -.15 (.12) | .20 (.17) |
| Age 3 Internalizing | -.18 (2.06) | .80 (.31)\*\* | -.09 (.16) |
| Age 3 Externalizing | .16 (.21) | .10 (.10) | .53 (.08)\*\*\* |
| Soc X Age 3 General factor | .00 (.02) | .01 (.01) | -.01 (.01) |
| Soc X Age 3 Internalizing | -.03 (.08) | -.06 (.05) | -.01 (.02) |
| Soc X Age 3 Externalizing | -.01 (.03) | -.01 (.01) | -.02 (.02) |
| **Maternal Parenting** |  |  |  |
| Authoritative | .00 (.00) | .00 (.00) | .00 (.00) |
| Age 3 General factor | .70 (.17)\*\*\* | -.26 (.37) | .31 (.41) |
| Age 3 Internalizing | -.03 (1.56) | 1.38 (1.61) | -.30 (1.35) |
| Age 3 Externalizing | .13 (.22) | .13 (.42) | .63 (.13)\*\*\* |
| Authoritative X Age 3 General factor | .02 (.01) | -.02 (.08) | 00 (.02) |
| Authoritative X Age 3 Internalizing | -.09 (.1) | .05 (.09) | -.01 (.04) |
| Authoritative X Age 3 Externalizing | .03 (.03) | .02 (.04) | -.01 (.02) |
| Authoritarian | 0 (.01) | 0 (.01) | .00 (.00) |
| Age 3 General factor | .72 (.41) | -.17 (1.23) | .36 (1.42) |
| Age 3 Internalizing | -.15 (5.43) | 1.37 (2.18) | -.27 (1.38) |
| Age 3 Externalizing | .21 (.55) | .13 (1.34) | .61 (.14)\*\*\* |
| Authoritarian X Age 3 General factor | -.02 (.05) | 0 (.07) | .01 (.06) |
| Authoritarian X Age 3 Internalizing | .02 (.4) | -.01 (.17) | 0 (.09) |
| Authoritarian X Age 3 Externalizing | -.04 (.07) | -.02 (.12) | -.01 (.05) |
| Permissive | .01 (.01) | .01 (.02) | -.01 (.01) |
| Age 3 General factor | .69 (.26)\*\* | -.19 (1.04) | .27 (4.98) |
| Age 3 Internalizing | -.17 (2.28) | .84 (4.31) | -.2 (.58) |
| Age 3 Externalizing | .19 (.3) | .02 (.64) | .62 (.14)\*\*\* |
| Permissive X Age 3 General factor | -.03 (.03) | -.06 (.28) | -.02 (.21) |
| Permissive X Age 3 Internalizing | .02 (.17) | .21 (.72) | 0 (.11) |
| Permissive X Age 3 Externalizing | -.06 (.05) | 0 (.11) | .02 (.07) |

\**p* < .05; \*\**p* < .01; \*\*\**p* < .001. Authoritative = Authoritative Parenting from the Parenting Styles and Dimensions Questionnaire; Authoritarian = Authoritarian Parenting from the Parenting Styles and Dimensions Questionnaire; Permissive = Permissive Parenting from the Parenting Styles and Dimensions Questionnaire.

Table S3. Differences in homotypic and heterotypic stability paths across levels of categorical moderators.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Age 6 General Factor | | | Age 6 Internalizing Factor | | | Age 6 Externalizing Factor | | |
|  |  | b (SE) | b (SE) | tMvF | b (SE) | b (SE) | tMvF | b (SE) | b (SE) | tMvF |
|  |  | Female | Male |  | Female | Male |  | Female | Male |  |
| Child Sex | Age 3 General | .64 (.15) | .80 (.23) | .81 | -.27 (.23) | -.36 (.35) | -.33 | .26 (.12) | .22 (.14) | -.24 |
|  | Age 3 Internalizing | .35 (.97) | -.70 (4.79) | -.25 | 1.42 (.57) | 2.05 (.80) | .86 | -.22 (.23) | -.15 (.21) | .27 |
|  | Age 3 Externalizing | .18 (.2) | .24 (.46) | .16 | .16 (.15) | .25 (.31) | .29 | .55 (.12) | .75 (.23) | .88 |
|  |  | No Dep | Dep | t+v- | No Dep | Dep | t+v- | No Dep | Dep | t+v- |
| Dep | Age 3 General | .63 (.17) | .72 (.26) | -.38 | -.27 (.67) | -.24 (.40) | -.08 | .2 (.12) | .29 (.24) | -.32 |
|  | Age 3 Internalizing | -.76 (2.81) | -.01 (2.50) | -.40 | 1.9 (1.43) | 1.6 (1.04) | .33 | -.37 (.44) | -.07 (.37) | -.80 |
|  | Age 3 Externalizing | .36 (.28) | -.09 (.24) | 1.50 | .17 (.21) | .06 (.32) | .28 | .61 (.15) | .61 (.15) | .04 |
|  |  | No Anx | Anx |  | No Anx | Anx |  | No Anx | Anx |  |
| Anx | Age 3 General | .72 (.19) | .65 (.25) | .29 | -.37 (.27) | -.26 (.31) | -.39 | .23 (.13) | .24 (.11) | -.06 |
|  | Age 3 Internalizing | -.32 (1.26) | .20 (3.66) | -.18 | 1.56 (.73) | 1.75 (.64) | -.26 | -.15 (.24) | -.17 (.25) | .06 |
|  | Age 3 Externalizing | .15 (.37) | .16 (.21) | -.02 | .34 (.26) | .03 (.11) | 1.37 | .64 (.18) | .53 (.13) | .55 |
|  |  | No SUD | SUD |  | No SUD | SUD |  | No SUD | SUD |  |
| SUD | Age 3 General | .64 (.15) | .73 (.20) | -.37 | -.30 (.24) | -.09 (.40) | -.54 | .30 (.14) | .15 (.13) | .90 |
|  | Age 3 Internalizing | .35 (.97) | -.01 (.76) | .68 | 2.16 (.94) | 1.22 (.74) | .77 | -.38 (.23) | -.08 (.22) | -1.08 |
|  | Age 3 Externalizing | .18 (.20) | .57 (.27) | -1.81 | .11 (.14) | .10 (.20) | .03 | .36 (.12) | .87 (.17) | -2.76\*\* |

Dep = Lifetime Maternal or Paternal Depressive Disorder; Anx = Lifetime Maternal or Paternal Anxiety Disorder; SUD = Lifetime Maternal or Paternal Substance Use Disorder; t+v- = t-statistic for disorder present vs. disorder absent

Figure S1. Moderation of the stability of the externalizing factor