Promises and Pitfalls of Latent Variable Approaches to Understanding Psychopathology: Reply to Burke and Johnston, Eid, Junghänel and colleagues, and Willoughby

G. Leonard Burns,
Washington State University

Christian Geiser,
Utah State University

Mateu Servera,
University of the Balearic Islands

Stephen P. Becker,
Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine

Theodore P. Beauchaine
The Ohio State University

Abstract

The commentaries by Burke and Johnston (this issue), Eid (this issue), Junghänel et al. (this issue), and Willoughby (this issue) on Burns et al. (this issue) provide useful context for comparing three latent variable modeling approaches to understanding psychopathology—the correlated first-order syndrome-specific factors model, the bifactor $S-1$ model, and the symmetrical bifactor model. The correlated first-order syndrome-specific factors model has proven useful in constructing explanatory models of psychopathology. The bifactor $S-1$ model is also useful for examining the latent structure of psychopathology, especially in contexts with clear theoretical predictions. Joint use of correlated first-order syndrome-specific model and bifactor $S-1$ model provides leverage for explaining psychopathology, and both models can also guide individual clinical assessment. In this reply, we further clarify reasons why the symmetrical bifactor model should *not* be used to study the latent structure of psychopathology and also discuss a restricted bifactor $S-1$ model that is equivalent to the first-order syndrome-specific factors model.

Keywords

ADHD; bifactor model; restricted bifactor $S-1$ model; factor analysis; p factor; attention-deficit/hyperactivity disorder; oppositional defiant disorder; psychopathology

Address correspondence to: G. Leonard Burns, Department of Psychology, Washington State University, Pullman, WA 99164-4820 glburns@wsu.edu.

Latent variable models play an important role in understanding psychopathology. Beginning with early work in the mid-1960s to the present (Willoughby, this issue), the correlated first-order syndrome-specific factors model has been instrumental in explanations of psychopathology. This latent variable model has been used countless times to evaluate the internal validity of dimensions of psychopathology. Internal validity occurs with convergent and discriminant validity of symptoms and syndromes within and across methods, sources, occasions, and age groups.

It is a mistake, however, to consider a latent variable as a *cause* of covariation among symptoms in a syndrome (Preszler & Burns, 2019). Rather, latent variables are descriptive representations of common variance across sets of symptoms; additional research is required to establish causes and specify external correlates of a latent syndrome. Establishing causal linkages across levels of analyses (e.g., genetic, neural, environmental influences on latent symptom dimensions) is, however, a complex endeavor (Beauchaine & Hinshaw, in press). With this context in mind, we discuss the four commentaries to Burns et al. (this issue).

Commentaries

As Eid (this issue) elegantly describes, the correlated first-order syndrome-specific factors model (Figure 1A) and bifactor $S-1$ model (Figure 1B) are *both* important in studies of psychopathology. First, the correlated first-order syndrome-specific factors model is essential for developing measures of psychopathology. Second, this model is essential for establishing convergent and discriminant validity of scores from such measures. Third, this model is used in many studies to determine external correlates of latent symptom dimensions. Thus, the correlated factors model has a clear role in descriptive and explanatory research on psychopathology.

After establishing validity of syndromes with the correlated factors model, the bifactor $S-1$ model can be used for more theory-driven studies of the structure of psychopathology (Eid, this issue; Heinrich et al. 2018). Joint use of both models is not to determine the better model, but rather to provide a richer context for understanding specific predictions from the bifactor $S-1$ model. As also mentioned by Eid (this issue), the first-order correlated factors model and a restricted version of the bifactor $S-1$ model can be formulated as equivalent models (see our discussion below). Eid's guidelines for use of both models should prove useful for research and clinical (individual assessment) activities.

Junghänel and colleagues (this issue) replicate our findings in the use of the bifactor $S-1$ model with attention-deficit/hyperactivity disorder (ADHD)/oppositional defiant disorder (ODD) symptoms among children and adolescents with ADHD/ODD diagnoses. It was encouraging to see such consistent findings across the studies. Such similar results speak well for use of psychometrically sound models to advance knowledge of psychopathology.

Burke and Johnston (this issue) and Willoughby (this issue), however, raise important issues about the symmetrical bifactor model and bifactor $S-1$ model. Our commentary addresses these issues through discussion of six overlapping questions.

(1) Does the bifactor $S - 1$ model result in overfitting in a manner similar to the symmetrical bifactor model? The symmetrical bifactor model frequently shows better model fit than the correlated first-order factors model and has therefore been seen as the "preferred" model. Burke and Johnston (this issue) note that superior fit of the symmetrical bifactor model may in fact reflect overfitting (estimating more, unnecessary parameters). They raise the important question of whether the bifactor $S - 1$ model is similarly prone to overfitting.

It is true that the standard version of the bifactor $S - 1$ model estimates more parameters than the correlated first-order factors model (more loading parameters). As shown in Figure 1, for example, for a design with three facets and three indicators (observed variables such as items or specific symptoms), the correlated first-order factors model (Figure 1A) estimates $3 \times 3 = 9$ standardized factor loadings (λ). For the same design, the bifactor $S - 1$ model (Figure 1B) estimates $3 \times 3 = 9$ reference factor loadings γ (all 9 variables load onto the reference factor) plus $2 \times 3 = 6$ specific factor loadings δ (only the indicators pertaining to the two non-reference facets have specific factor loadings). Therefore, the standard bifactor $S - 1$ model estimates more loading parameters and may show better fit to the correlated factors model for a given set of data.

For cases in which overfitting is a concern, an alternative restricted bifactor $S - 1$ model can be used (Figure 1C and 1D) that has the same number of parameters and is equivalent to the correlated first-order factors model. To define this restricted bifactor $S - 1$ model, the correlated first-order factors model is used as a starting point. Instead of estimating the three correlations between the first-order factors F_f as shown in Figure 1A, the non-reference factors (here: F_2 and F_3) are regressed on the reference factor (here: F_1) using linear regression analysis (Figure 1C). Residuals of these regressions represent specific factors S_f. Specific factors in the $S - 1$ approach thus have very clear meaning: They represent variance specific to Facet f that cannot be linearly predicted from the reference facet.

In this model, all specific factors are allowed to correlate. The example shown in Figure 1C is equivalent to the model in Figure 1A. The structural (latent variable) model in Figure 1C has the same number of parameters as the structural model in Figure 1A. Instead of estimating three latent correlations, we now estimate two regression slope coefficients (β_f) and one latent residual correlation between the two specific factors S_1 and S_2.

A mathematical reformulation of this latent regression model yields the equivalent restricted bifactor $S - 1$ model shown in Figure 1D.[1] In this model, indicators pertaining to the reference facet ($f = 1$) have the exact same loadings (λ) as in the correlated first-order factors model (Figure 1A). Thus, the general reference factor in the restricted bifactor $S - 1$ model is identical to the factor F_1 in the correlated first-order factors model. General reference factor loading parameters of the non-reference indicators are constrained to equal

---

[1]This can be seen by writing out the measurement and structural equations in the model shown in Figure 1C. The measurement equation for mean-centered variables is $Y_{if} = \lambda_{if} F_f + \varepsilon_{if}$ (i = indicator, f = facet). The structural (latent regression) equation for the non-reference facets is $F_f = \beta_f F_1 + S_f$, where $f = 1$ indicates the reference facet. Inserting the structural equation for F_f into the measurement equation for the non-reference facets yields $Y_{if} = \lambda_{if}(\beta_f F_1 + S_f) + \varepsilon_{if} = \lambda_{if}\beta_f F_1 + \lambda_{if}S_f + \varepsilon_{if}$. This is the measurement equation for the restricted bifactor $S - 1$ model shown in Figure 1D.

the products of the relevant λ and β parameters. The number of independent loading parameters in this model is 11 (9 independent λ and 2 independent β parameters). In addition, there is only 1 latent correlation (between S_1 and S_2) rather than 3 latent correlations as in the correlated first-order factors model (which estimates correlations between all three F_f factors). Therefore, the model in Figure 1D has the same total number of parameters as the correlated first-order factors model in Figure 1A.

For a given set of data, the restricted bifactor $S-1$ model yields the exact same fit as the correlated first-order factors model and thus there is no overfitting relative to the correlated first-order factors model. A researcher can thus use the restricted bifactor $S-1$ model if there is a concern with overfitting with the bifactor $S-1$ model. The conceptual meaning of the general reference and specific factors in the restricted bifactor $S-1$ model is the same as in the bifactor $S-1$ model (Geiser, Eid, & Nussbeck, 2008).

Appendix A shows Mplus code, global fit indices, factor loadings, and factor variances/correlations for the four models with our data. Figure 1A, 1C, and 1D are shown to be equivalent models. These results also show the meaning of the general reference and specific factors is the same across the bifactor $S-1$ and restricted bifactor S-1 models. ADHD-HI symptom loadings are the same across the four models as well.

(2) Does selection of a different general reference factor in the bifactor $S-1$ model change findings? Burke and Johnston (this issue) also raise the important question of whether selection of a different reference facet changes the interpretation of results from the bifactor $S-1$ approach. The answer is yes, (and this change is expected). To explain, we use an analogy to regression analysis with dummy-coded nominal independent variables. In the same way as interpretation of regression coefficients (intercepts and slopes) changes when a different reference group is selected in dummy coding, interpretation of reference and specific factors changes in the $S-1$ approach when a different symptom group or facet is selected as reference. This can be seen most easily from Figures 1C and 1D, which depict the restricted bifactor $S-1$ model. When a different first-order factor (e.g., F_2 instead of F_1) is selected as reference and the remaining factors are regressed on this factor, the meaning of the general reference factor changes to equal the meaning of the different first-order factor (e.g., F_2). It is therefore important to carefully consider *a priori* which facet should serve as reference. This choice should be based on theoretical grounds and/or ease of interpretation.

Note that when using the restricted bifactor $S-1$ approach, overall model fit remains the same regardless of which facet (F_1, F_2, or F_3) is used as reference facet (Geiser et al., 2008). This is again analogous to dummy coding in regression analysis where the overall R^2 value is unaffected by the choice of the reference category in dummy coding.

(3) Why does the general factor in the symmetrical bifactor model change its meaning from a general to a specific factor when one specific factor collapses or when one or more indicators have specific factor loadings of zero? Willoughby (this issue) questioned one of our central arguments—that is, *the meaning of the general factor in the symmetrical bifactor model changes from a general to a specific factor when one specific factor collapses or when one or more indicators have zero loadings on their specific factors.* Specifically,

Willoughby (this issue) commented that "Burns and colleagues (…) maintain that [in this case] the general factor should be interpreted as the HI factor. I do not understand how this can be true." (p. xx). Willoughby argued that all 18 items still contribute to the general factor given that they all load onto this factor. We believe that this confusion is shared by many researchers. We therefore elaborate in more detail immediately below.

To understand why the general factor turns into a specific factor under the conditions described above (factor collapse or non-significant individual specific factor loadings), reconsider our discussion of the restricted bifactor $S-1$ model. Figures 1C and 1D show that the restricted bifactor $S-1$ model is a direct mathematical reformulation of the correlated first-order factors model (Figure 1A). As we explained above, the models in Figures 1A, 1C, and 1D are equivalent and will always fit equally well for a given data set (Appendix A).

Most researchers would agree that the factors in the correlated first-order factors model (Figure 1A) are facet-specific factors (e.g., HI, IN, and ODD factors). Reformulation of this model into an equivalent restricted bifactor $S-1$ model (Figures 1C and 1D) shows us that the reference factor F_1 remains *the exact same facet-specific factor* (e.g., an HI factor), even though all items now load onto this reference factor as shown in Figure 1D. In other words, the fact that all items load onto a factor does not necessarily make this factor a general factor. The same issue occurs when a specific factor collapses empirically in the symmetrical bifactor model, leading to a reduction of this model to an $S-1$ model and a change in the interpretation of the general factor to a facet-specific factor.[2] Table S2 shows the ADHD-HI symptom loadings remain the same across the four models (1A, 1B, 1C, and 1D).

(4) Can the bifactor $S-1$ model be used to test additional hypotheses about the structure of ADHD and ODD symptoms? We describe one additional theoretical-driven hypothetical study to address this question. Other theoretical-driven studies are also possible; though space only allows one example.

ODD symptoms are considered to involve separate irritability and defiant factors (Burke & Johnston, this issue). To date, the correlated first-order syndrome-specific factors model and the symmetrical bifactor model have been used to evaluate the internal and external validity of the two ODD factors (Burke and Johnston, this issue). The bifactor $S-1$ model could be used to further evaluate distinctiveness of the two ODD factors relative to the correlated factors model.

One possibility would be to use the ADHD-HI symptoms to define the general reference factor, with specific ODD irritability and ODD defiant residual factors. Given evidence that ADHD-HI symptoms confer considerable vulnerability to later development of ODD, whereas ADHD-IN symptoms do not (Beauchaine, Zisner, & Sauder, 2017), using the bifactor $S-1$ model would allow for examination of true score variance in ODD irritability

---

[2]Note that the fact that factors have the same loadings does not always mean that the factors have the same meaning. This can be seen by comparing the factors F_f and S_f in Figures 1A and 1D. Notice that F_2 and F_3 in Figure 1A have the same loadings as do S_2 and S_3 in Figure 1D. Nonetheless, the factors F_f and S_f have different meanings. The F_f factors are true score variables, whereas the S_f factors are true score regression residual variables. This is because the indicators measuring the S_f factors *also* load on the reference factor F_1 in Figure 1D.

and ODD defiant factors independent of the general ADHD-HI reference factor as well as discriminant validity of the two ODD residual factors. A developmental psychopathology hypothesis could be that the amount of residual true score variance in the two ODD specific factors (after accounting for shared variance with the general ADHD-HI factor) increases with age. Finally, the model could determine if the two specific ODD residual factors have common and unique external correlates, independent of the general reference factor. Such a model would yield clear, psychometrically sound definitions of the two specific ODD residual factors.

(5) Are problems associated with the symmetrical bifactor model restricted to symptom ratings (item-level data)? Willoughby (this issue) suggest that problematic issues with the symmetrical bifactor model that we highlighted in Burns et al. (this issue) are only an issue for item-level data: "It is important to appreciate that the problems that Burns and colleagues raised with respect to B_{Sym} [the symmetrical bifactor model] were specific to questions that are focused on the structure of *item-level* data" (p. xx; emphasis in the original). We disagree as we believe this issue is much more general in nature. It appears to cause problems whenever researchers apply the symmetrical bifactor model to structurally different (as opposed to exchangeable) facets. Heinrich et al. (2018) showed that the same issues occurred in attempts to identify a general factor of depression. Studies of the *p* factor of psychopathology (e.g., Caspi et al. 2014) typically use scale-level data and are prone to the same issues that we highlighted for ADHD/ODD symptom ratings. For example, Caspi et al. (2014) examined three first-order factors (internalizing, externalizing, thought disorder) and reported a negative specific factor variance in their bifactor model. According to Caspi et al., this indicated that no specific factor could be identified for the thought disorder facet. Caspi et al. subsequently dropped the specific factor for thought disorder from their model. Thus, the general factor became a specific thought disorder factor. The same issues occurred in bifactor applications to summary scores of intelligence facets (Eid et al. 2018) as well as in multitrait-multimethod studies in which structurally different raters (e.g., mothers and teachers) serve as "facets" (Eid et al., 2008).

In summary, problems with the symmetrical bifactor model are not limited to item-level (symptom) data. Similar problems result any time the symmetrical bifactor model is used with structurally different facets or raters. Thus, it does *not* matter whether the manifest variables are individual symptoms or summary (interval-scale) scores on a symptom dimension such as in *p* studies—the statistical issues remain the same, and we thus discourage use of the symmetrical bifactor model not just for item-level (symptom) data, but more generally whenever researchers analyze structurally different facets or data from non-interchangeable raters.

(6) Can we obtain a general factor that represents a general propensity to psychopathology as originally intended by users of the symmetrical bifactor approach? Maybe one of the most important criticisms from Willoughby (this issue) concerns an original goal of researchers who apply bifactor models: to identify a *general* overarching factor that represents "general psychopathology." As Willoughby correctly points out, such a general factor is *not* obtained in the $S-1$ model unless this general factor is directly measured by a separate set of indicators as proposed by Eid (this issue). Therefore, as

mentioned by Willoughby as well as Burke and Johnston (both in this issue), some questions addressed by the $S-1$ approach differ from the answers sought with the symmetrical bifactor model.

Although we understand the desire to identify a general psychopathology dimension, we argue that the symmetrical bifactor approach does not hold this promise. This may be a sign that different ways of identifying a "general factor of psychopathology" should be developed and used, and/or that related questions should be asked in different ways. As we show in our original article and explain in more detail in this reply, the originally-intended general psychopathology factor is typically *not* obtained in empirical applications of the symmetrical bifactor model. As we demonstrate above, the symmetrical bifactor model reduces to an $S-1$ model (with a specific factor defining the general factor) when one specific factor collapses or when one or more variables do not load significantly on their specific factors. As we show in Burns et al. (this issue), this happens in the majority of applications to ADHD symptom reports, and so the intended general factor is not actually found as also occurs in p studies (e.g., Caspi et al. 2014).[3]

Conclusions

Latent variable models remain quite useful—and perhaps essential—to our understanding of psychopathology. The correlated first-order syndrome-specific factors model has proved especially useful for development of measures of psychopathology. This model has also proved useful for identifying external correlates of latent symptom dimensions. These two uses—establishment of internal and external validity—are critical toward explaining rather than describing psychopathology. Individual clinical assessments benefit from this model as well.

Although the symmetrical bifactor model enjoys widespread appeal, this model should *not* be used to examine the latent structure of psychopathology. *Use of this model requires randomly selected and thereby interchangeable facets, whereas symptom dimensions represent structurally different facets.* Violation of this assumption usually yields inadmissible/anomalous solutions that often result in the general factor becoming facet- or even symptom-specific. In contrast, the bifactor $S-1$ model is fully appropriate to use when examining the latent structure of psychopathology, especially when used with clear theoretical predictions, and in conjunction with the correlated first-order syndrome-specific factors model. Together, the correlated first-order syndrome-specific factors model and the bifactor $S-1$ model hold considerable potential to advance our understanding of psychopathology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

---

[3]The application of the symmetrical bifactor model to structurally different symptom facets can yield an admissible solution if the correlations among the symptom indicators are homogenous, the patterns of correlations common with randomly selected and thereby inter-changeable raters. Such an admissible solution, however, still does not yield a psychometrically sound definition of the general factor (see Eid et al. 2017; Heinrich et al. 2018).

Acknowledgments

References

Beauchaine TP, & Hinshaw SP (in press). The Research Domain Criteria and psychopathology among youth: Misplaced assumptions and an agenda for future research. Journal of Clinical Child and Adolescent Psychology. doi:10.1080/15374416.2020.1750022

Beauchaine TP, Zisner A, & Sauder CL (2017). Trait impulsivity and the externalizing spectrum. Annual Review of Clinical Psychology, 13, 343–368. doi:10.1146/annurev-clinpsy-021815-093253

Burke JD, & Johnston OG (this issue). The bifactor S - 1 model: A psychometrically sounder alternative to test the structure of ADHD and ODD? Journal of Abnormal Child Psychology. doi:10.1007/s10802-020-00645-4

Burns GL, Geiser C, Servera M, Becker SP, & Beauchaine TP (this issue). Application of the bifactor $S-1$ model to ADHD/ODD symptoms: An appropriate model bifactor model for symptom ratings. Journal of Abnormal Child Psychology. doi:10.1007/s10802-019-00608-4

Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, …, Moffitt TE. (2014). The p Factor: One general psychopathology factor in the structure of psychiatric disorders. Clinical Psychological Science, 2, 119–137. doi:10.1177/2167702613497473 [PubMed: 25360393]

Eid M (this issue). Multi-faceted constructs in abnormal psychology: Implications of the bifactor $S-1$ model for individual clinical assessment. Journal of Abnormal Child Psychology. doi: 10.1007/s10802-020-00624-9

Eid M, Geiser C, Koch T, & Heene M (2017). Anomalous results in G-factor models: Explanations and alternatives. Psychological Methods, 22, 541–562. doi:10.1037/met0000083 [PubMed: 27732052]

Eid M, Krumm S, Koch T, & Schulze J (2018). Bifactor models for predicting criteria by general and specific factors: problems of nonidentifiability and alternative solutions. Journal of Intelligence, 6(3): 42. doi: 10.3390/jintelligence6030042

Eid M, Nussbeck FW, Geiser C, Cole DA, Gollwitzer M, & Lischetzke T (2008). Structural equation modeling of multitrait-multimethod data: Different models for different types of methods. Psychological Methods, 13, 230–253. doi:10.1037/a0013219 [PubMed: 18778153]

Geiser C, Eid M, & Nussbeck FW (2008). On the meaning of the latent variables in the CT-C(M–1) model: A comment on Maydeu-Olivares and Coffman (2006). Psychological Methods, 13, 49–57. doi:10.1037/1082-989X.13.1.49 [PubMed: 18331153]

Heinrich M, Zagorscak P, Eid M, & Knaevelsrud C (2018). Giving G a meaning: An application of the bifactor-(S-1) approach to realize a more symptom-oriented modeling of the Beck Depression Inventory–II. Assessment. Advance online publication. doi:10.1177/1073191118803738.2016.1144192

Junghänel M, Rodenacker K, Dose C, & Döpfner M (this issue). Applying the bifactor $S-1$ model to ratings of ADHD/ODD Symptoms: A commentary on Burns et al. (2019) and a re-analysis. Journal of Abnormal Child Psychology. doi: 10.1007/s10802-020-00637-4.

Preszler J, & Burns GL (2019). Network analysis of ADHD and ODD symptoms: Novel insights or redundant findings with the latent variable model? Journal of Abnormal Child Psychology, 47, 1599–1610. doi: 10.1007/s10802-019-00549-y [PubMed: 31025233]

Willoughby MT (this issue). Commentary on the application of the bifactor $S-1$ model to multisource ratings of ADHD/ODD symptoms: An appropriate bifactor model for symptom ratings. Journal of Abnormal Child Psychology. doi:10.1007/s10802-020-00629-4
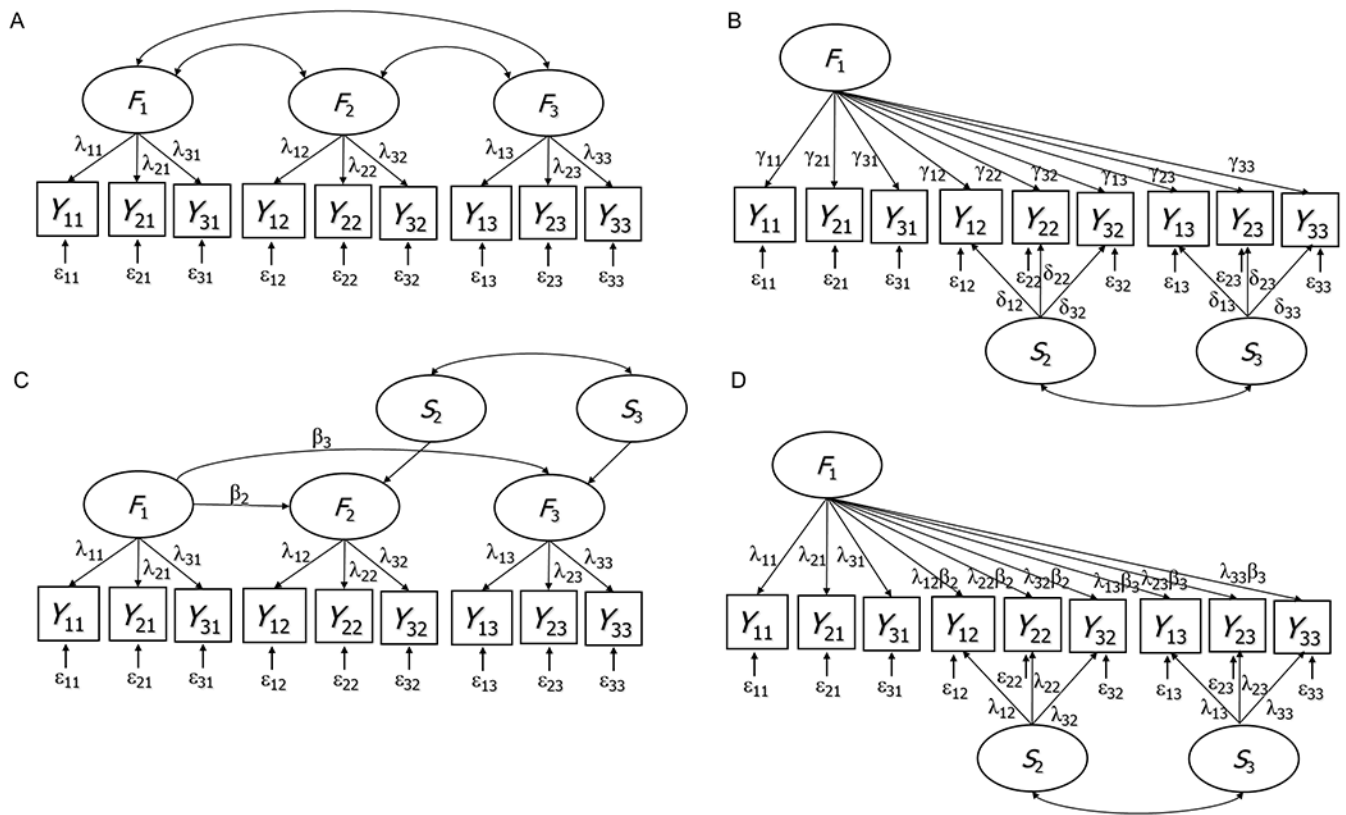
Figure 1.
Confirmatory factor models for multi-facet data. Y_{if}: ith observed variable or symptom pertaining to facet (or syndrome) f. F_f: true score variable (latent factor) pertaining to facet f. S_f: specific (residual) factor pertaining to facet f. λ_{if}, γ_{if}, δ_{if}: factor loadings. β_f: latent regression slope coefficient. A: correlated first-order factors model. B: Standard bifactor $S - 1$ model. C: Reformulation of the correlated first-order factors model into a latent regression model. D: Reformulation of the latent regression model into a restricted bifactor $S - 1$ model. Models A, C, and D are equivalent. In Models B, C, and D, Facet 1 serves as reference. Note that in Model D, the reference factor loadings of the non-reference indicators are constrained to equal the product $\lambda_{if} \cdot \beta_f$.