In the absence of a gold standard

Noordhof, Arjen

Chapter 4

Optimal use of multi-informant data on co-occurrence of internalizing and externalizing problems. *

Arjen Noordhof, Albertine J. Oldehinkel, Frank C. Verhulst, Johan Ormel

Abstract

Strong between-informant discrepancies are found in ratings of (pre)adolescent problems and in co-occurrence rates between different domains of psychopathology. These discrepancies can be caused by differences in the context of measurement and the perspective of informants (Kraemer et al., 2003). The aim of this study was to develop a 'Multi-Informant Co-occurrence' model (MIC), which takes into account these differences in context and perspective. In a population-based cohort of (pre)adolescents (n=2230) from a longitudinal study in the North of the Netherlands, internalizing (INT) and externalizing (EXT) problems were rated by the (pre)adolescents themselves, their teachers, and their parents. As hypothesized Principal Component Analysis revealed four independent main components: Between-domain convergence was captured by a severity component (S), while between-domain discrepancy was captured by a direction component (D). Between-informant discrepancies were captured by a perspective (P) and a context (C) component. The use of this MIC-model will increase reliability and validity of measures of psychopathology and the four components each provide useful specific information.

Introduction

There is no gold standard for the measurement of psychiatric symptoms and disorders, and reports from different informants tend to correlate only moderately (Achenbach, et al., 1987). Therefore, in most circumstances using multiple informants seems the best strategy to chart mental health problems (Offord, et al., 1996), and there is a need for theory-based approaches to combine their reports (De Los Reyes & Kazdin, 2005). Based on a testable theory about why and how informants converge and diverge, Kraemer and colleagues (2003) have proposed a pragmatic approach for combining information from multiple sources, involving a clear and applicable research design, and a straightforward statistical procedure. This method has been applied successfully to reports of psychopathology (Kraemer, et al., 2003).

While Kraemer and colleagues (2003) applied their model to internalizing and externalizing problems separately, scores of these two problem domains typically correlate substantially (Lilienfeld, 2003). Such correlations can be related to

* *Published as an article in 2008 in The International Journal of Methods in Psychiatric Research, vol. 17(3), pg. 174-183.*

comorbidity of underlying disorders, but might be influenced by informant biases as well[6]. Co-occurrence rates tend to diverge strongly between informants (Keiley, Bates, Dodge, & Pettit, 2000), and different procedures of combining multi-informant information can lead to very different estimates of co-occurrence (Youngstrom, Findling, & Calabrese, 2003). Furthermore, Burt and colleagues (2005) showed that heritability rates of co-occurrence (of various externalizing disorders) depend on the way data from multiple informants are combined. Thus, the issue of convergence and divergence between informants is not only relevant within each domain, but also in understanding the relation between the domains of internalizing and externalizing problems. In the present study we extended Kraemer and colleagues' model (2003) to analyze the co-occurrence of problems from both the internalizing and the externalizing domain in a general population sample of pre- and early adolescents.

Central to the approach of Kraemer and colleagues (2003) is the idea that the reports of informants are not simply measurements of a well-known characteristic (G), but are influenced by the context (C) of observation and the perspective (P) of a specific informant Thus, the score of an informant is not just the result of the characteristic and measurement error, but is also influenced by context and perspective. Taking context, perspective, and measurement error into account may increase the reliability and validity of the estimated mental health problems, and provide information on the role of these aspects.

In (pre)adolescents two important contexts (C) are school and home, and two important perspectives (P) self-report (from inside) and other-report (from outside). To obtain information on these components, it is essential to carefully select informants that are likely to diverge strongly on one aspect, but converge on the other. For example, teachers and parents are likely to diverge strongly regarding the context, but both represent an 'outside'-perspective. Youngsters themselves on the other hand are likely to diverge most strongly from both their parents and teachers based on the difference in perspective. As Kraemer and colleagues (2003) substantiated in their study, the reports from teachers, parents and the (pre)adolescents themselves are well-suited to estimate these C and P components independently from the general characteristic (G). Using Principal Component Analysis (PCA) for reports of these three informants in three different samples, they found support for the hypothesis that: [1] the informants had similar coefficients on the general component (G); [2] parents and teachers had similar coefficients on the perspective component (P), both reverse to the self-report coefficients; and [3] parents and teachers had opposite coefficients on the context component (C), while the self-reports were in between.

6 As has been discussed by Lilienfeld (2003) and Meehl (2001), the term comorbidity is meaningful only in the context of well-validated disease entities. As still little is known about diseases underlying psychiatric classifications, we prefer using the term co-occurrence as indicating that a person can be classified in more than one psychiatric category.

The orthogonality of these components (i.e. G, P and C do not correlate) follows from their definition. Perspective is defined as the variance that is only due to the difference between self and others. Context is defined as the variance that is only due to the difference between ratings at school and at home. As the teacher- and parent-ratings diverge on C, while they converge on P, these components are defined as orthogonal. G is defined as the convergence of all informants, which is orthogonal to C and P, because C and P are defined by divergence of two of the informants.

The aim of the present study was to extend this model, in order to capture co-occurrence of problems from the internalizing (INT) and externalizing (EXT) domain. Co-occurrences of different psychiatric categories can be captured using a hierarchical model (Krueger, 1999; Krueger, et al., 2003; Vollebergh, et al., 2001). Therefore, as shown by Weiss, Susser and Catron (1998), disorders can be described as combinations of *narrow-band* features, differentiating specific categories (e.g. depression, anxiety), *broad-band* features, differentiating internalizing and externalizing disorders, and *common* features, related to the general severity of disorders. A similar approach is to discern between a severity component (S) and a direction component (D), which correspond to the common and broad-band specific aspects respectively (Essex, et al., 2006; Ormel, et al., 2005). These S- and D-scores are by definition uncorrelated and easily computable (Essex, Klein, Cho, & Kraemer, 2003). The strength of that approach is that it captures both the co-occurrence of disorders and their differentiation. We expected that combining this approach with the model by Kraemer et al. (2003) would result in a useful tool for analyzing multi-informant measurements of co-occurring psychiatric problems.

We used the reports of different informants on a number of specific scales from both the INT and EXT domain. Instead of a different model for each informant or for each specific disorder, one model was developed that would explain the covariances between all these scores. It was hypothesized that [1] between-informant discrepancies can be explained by the C- and P-components, which should have the same patterns of coefficients as those found by Kraemer and colleagues (2003); [2] between-domain discrepancies can be explained by broad-band features, resulting in a 'direction'-component (D) on which problem-scales from the INT and EXT domain have opposite coefficients; and [3] between-domain convergence can be explained by common features, resulting in a 'severity'-component (S) on which problem scales from the INT and EXT domain have equal coefficients. Thus, to explain both between-informant and between-domain covariance, at least four components (C, P, D and S) are necessary. In principle, however, the model does not exclude additional components; because context (C) and perspective (P) might be related to common features, broad-band features, or both; and because the narrow-band features might influence the ratings as well. Thus, the analysis was aimed at testing the hypothesis that at least four components (C, P, D and S) would be found, and at exploring the possibility that more components would emerge. Furthermore, we aimed at developing a method by which these components can easily be applied in research and clinical practice.

Methods

Sample

Subjects were participants of the 'Tracking Adolescents' Individual Lives Survey' (TRAILS), a prospective cohort study of Dutch (pre)adolescents. The present study involves data from the first (T1) and second (T2) assessment wave of TRAILS, which ran from March 2001 to July 2002 and September 2003 to December 2004, respectively. A detailed description of the sampling procedure and methods is provided by De Winter and colleagues (De Winter, et al., 2005).

Briefly, the TRAILS target sample involved all 10- to 11-year-old children living in the three largest cities and some rural areas in the North of The Netherlands. Of the eligible children, 76.0% (n=2230) were enrolled in the study. Responders and non-responders did not differ regarding the prevalence of teacher-rated problem behavior and associations between sociodemographic variables and mental health indicators (de Winter et al., 2005).

Of the 2230 baseline (T1) participants, 96.4% (n=2149, 51.2% girls) participated in the first follow-up assessment (T2), which was held 2–3 years after T1 (mean number of months 29.44, S.D.=5.37). Mean age at T2 was 13.55 (S.D.=0.54). After complete description of the study to the subjects, written informed consent was obtained from both the (pre)adolescent and the parents.

Instruments

At T1 and T2, questionnaires were completed by teachers, parents and the (pre)adolescents themselves. The parent-rated Child Behavior Checklist (CBCL; Achenbach, 1991a), the Youth Self Report (YSR; Achenbach, 1991c) and the Teacher Checklist of Psychopathology (TCP; De Winter et al., 2005) were used to assess psychopathology. The CBCL and YSR have been developed for the multi-informant assessment of child and adolescent psychopathology. In these 112-item questionnaires, the informant rates descriptions of emotions and behaviors as not [0], sometimes [1], or very often [2] present. Factor analysis on these items revealed a structure of eight syndrome scales (De Groot, Koot, & Verhulst, 1994). Three of these were related to the internalizing domain (INT): 'Anxious-depressed' (Anx), 'Somatic complaints' (Sc), and 'Withdrawn/depressed' (Wd). Two were related to the externalizing domain (EXT): 'Aggressive behavior' (Agg) and 'Rule-Breaking behavior' (Rb). The TCP contains descriptions (vignettes) of problem behaviors, corresponding to the eight syndromes of the CBCL and YSR. Teachers rated each of these vignettes on a 5-point rating scale. At T1, for 79% of the participants these questionnaires were completed by all three informants, and for 97% by at least two informants. At T2, for 60% of the original (T1) participants, questionnaires were completed by all informants, and for 91% by at least two informants. The low three-informant rates at T2 were mainly caused by a relatively low response rate among teachers. Only children for which three informants completed the questionnaires were included in the analyses.

Statistical analysis

All analyses were performed using *SPSS 12.0.2*. For each of the five syndrome scales related to the INT and EXT domain, we calculated a mean item score (i.e. five scores for each of the three informants). Principal Component Analysis (PCA) with these 15 (3x5) subscales was used to test the hypotheses. As explained in the introduction the components are defined as orthogonal and are hypothesized to explain a large part of the variance in the subscales. For these specific hypotheses Principal Component Analysis (PCA) is an appropriate tool, as it is a method that maximizes the explained variance of each orthogonal component.

Following the approach of Kraemer and colleagues (2003) we used the unrotated principal components. Subsequently, the stability of the model across different ages was investigated, by performing the same analyses on the T2-data.

Table 1. *Contrasts used for the interpretation of the predicted PCA-components.*

	Interpretation of coefficients				
Severity		General (**G**)			
Direction		Internalizing (**I**)	*vs.*	Externalizing (**E**)	
Perspective		Self (**Se**)	*vs.*	Others (**Ot**)	
Context	Home (**Ho**) *vs.*	School (**Sc**)	*vs.*	Both (**B**)	

Table 1 shows the four different components that we expected based on these hypotheses, and for each of these components the expected contrasts. For all components we evaluated whether they matched the hypothesized pattern of: either [1] a severity component (S), indicated by similar coefficients of the subscales (table 1: G); [2] a direction component (D), indicated by a contrast between coefficients of INT and EXT subscales (table 1: I vs E); [3] a perspective component (P), indicated by a contrast between self-rated and other-rated problems (table 1: Se vs Ot); or [4] a context component (C), indicated by a contrast between school and home (table 1: Ho vs Sc, and B for involvement in both contexts). All components that could be interpreted this way were included in the final model. Other components that were found in the analyses could not be interpreted a-priori. Such components can be either related to unique variance of a specific subscale, or to domain- or disorder-specific informant discrepancies. The number of possibly important components is high and post-hoc interpretations of these components can be rather arbitrary. To prevent irrelevant chance findings, other components were only interpreted (post-hoc) if their eigenvalue exceeded 1. The rationale for this, is that if the eigenvalue of a component does not exceed 1, that component contributes less explained variance than a single subscale in the original model (i.e., less than 7%). Variance that cannot be explained by our model is to be explained by specific subscales and specific biases that are certainly of interest, but require more specific hypotheses, which are beyond the scope of this article.

In order to enhance the applicability of the model for research and clinical practice we developed algorithms for S, D, P and C that can be made without performing a PCA. We followed the method used by Essex et al. (2003) to compute the S- and D-component, and extended it to the computation of the P- and C-component. We investigated to what extent individual scores based on these algorithms were comparable to scores based on the coefficients of the principal components by calculating individual component scores with the SCORES-option available in the SPSS factor analysis command (which can be used for PCA as well) and correlating these scores with the scores based on the simple algorithms. The severity component (S) was calculated by taking the mean of all 15 subscores; the direction component (D) by contrasting the internalizing and externalizing subscales; the perspective component (P) by contrasting ratings by others and self ratings; and the context component (C) by contrasting teacher and parent ratings. More specifically:

[1] S = mean (all subscores)
[2] D= [mean (internalizing subscales) – mean (externalizing subscales)] / 2
[3] P = [mean (parent and teacher subscales) – mean (self subscales)] / 2
[4] C= [mean (teacher subscales) – mean (parent subscales)] / 2

Results

Analyses were performed on data from the first (T1) and second (T2) assessment. Means and standard deviations of these data are reported in table 2.

Table 2. *Means and standard deviations for YSR, CBCL and TCP subscales at the first and second assessment wave.*

Informant	Subscale	data T1		data T2	
		Mean	*Sd*	Mean	*Sd*
Self	Anx	0.33	*0.27*	0.31	*0.29*
	Sc	0.43	*0.31*	0.33	*0.29*
	Wd	0.34	*0.29*	0.34	*0.30*
	Agg	0.31	*0.25*	0.31	*0.24*
	Rb	0.22	*0.17*	0.26	*0.20*
Parent	Anx	0.28	*0.25*	0.20	*0.22*
	Sc	0.20	*0.21*	0.16	*0.20*
	Wd	0.25	*0.27*	0.24	*0.28*
	Agg	0.35	*0.29*	0.23	*0.25*
	Rb	0.13	*0.13*	0.10	*0.14*
Teacher	Anx	0.69	*0.95*	0.78	*1.03*
	Sc	0.58	*0.86*	0.70	*0.96*
	Wd	0.72	*0.99*	0.98	*1.14*
	Agg	0.62	*0.98*	0.65	*1.10*
	Rb	0.27	*0.71*	0.32	*0.84*

Note: *Anx* = Anxiety; *Sc* = Somatic complaints; *Wd* = Withdrawn behavior; *Agg* = Aggression; *Rb* = Rule-Breaking behavior.

Analysis of T1-data

The result of the PCA are reported in table 3. Four components were found that had eigenvalues greater than 1 and followed the hypothesized between-informant and between-domain convergences and divergences. Other components did not meet the criterion of an eigenvalue greater than 1, and therefore were not included in the final model. Therefore, four components were included in the final model, which explained 64% of the total variance.

Two components were found with a rather high between-informant convergence. These were interpreted as a severity component (S), with positive coefficients for all subscales, and a direction component (D), with a contrast between the coefficients of the INT (+; positive coefficients) and the EXT (-; negative coefficients) domain. The other two components matched the hypothesized between-informant divergences: a perspective component (P), with a strong contrast between the coefficients of the self-rated (-) and the other-rated (+) subscales; and a context component (C), with a strong contrast between the coefficients of the parent-rated (-) and the teacher-rated subscales (+), and with the self-rated subscales in between (only weak coefficients <.20).

Table 3. *Component coefficients for the joint analysis of the internalizing and externalizing domain.*

Informant	Subscale	**S**everity		**P**erspective		**D**irection		**C**ontext	
Self	Anx	0.58	**G**	-0.51	**Se**	0.30	**I**	0.16	**B**
	Sc	0.46	**G**	-0.49	**Se**	0.15	**I**	0.15	**B**
	Wd	0.58	**G**	-0.45	**Se**	0.28	**I**	0.19	**B**
	Agg	0.62	**G**	-0.51	**Se**	-0.26	**E**	0.06	**B**
	Rb	0.52	**G**	-0.47	**Se**	-0.35	**E**	0.05	**B**
Parent	Anx	0.59	**G**	0.23	**Ot**	0.35	**I**	-0.41	**Ho**
	Sc	0.44	**G**	0.12	**Ot**	0.30	**I**	-0.34	**Ho**
	Wd	0.58	**G**	0.28	**Ot**	0.31	**I**	-0.32	**Ho**
	Agg	0.67	**G**	0.27	**Ot**	-0.24	**E**	-0.42	**Ho**
	Rb	0.61	**G**	0.25	**Ot**	-0.33	**E**	-0.41	**Ho**
Teacher	Anx	0.42	**G**	0.45	**Ot**	0.22	**I**	0.50	**Sc**
	Sc	0.37	**G**	0.35	**Ot**	0.20	**I**	0.48	**Sc**
	Wd	0.32	**G**	0.43	**Ot**	0.33	**I**	0.47	**Sc**
	Agg	0.43	**G**	0.27	**Ot**	-0.61	**E**	0.30	**Sc**
	Rb	0.37	**G**	0.24	**Ot**	-0.61	**E**	0.27	**Sc**
% Explained variance		**27%**		**14%**		**12%**		**11%**	

Note: *Anx* = Anxiety; *Sc* = Somatic complaints; *Wd* = Withdrawn behavior; *Agg* = Aggression; *Rb* = Rule-Breaking behavior, *G*=general; *Se*=self; *Ot*=other; *I*=internalizing; *E*=externalizing; *Ho*=Home; *Sc*=School; *B*=both.

Replication at T2

Results of the PCA on T2-data are shown in table 4. The same components were found, with only small differences in component-coefficients. One exception was that the two teacher-rated externalizing subscales (Agg and Rb) loaded less strongly on the P-component than at T1. Only minor changes in explained variance of the four components occurred. This model explained 64% of variance, which is exactly the same as the model based on T1-data. The solution for the D-component was opposite to the solution found with the T1-data, as the externalizing scales loaded positive and the internalizing scales loaded negative. However by multiplying all scores with -1, which is statistically justified, it is possible to obtain a completely similar solution, that is with high D-scores indicating relatively more internalizing problems.

Table 4. *Component coefficients for the joint analysis of the internalizing and externalizing domain at T2.*

Informant	Subscale	Severity		Direction		Perspective		Context	
	Anx	0.54	**G**[2]	-0.49	**I**	-0.46	**Se**	0.09	**B**
	Sc	0.48	**G**	-0.32	**I**	-0.45	**Se**	0.07	**B**
Self	Wd	0.52	**G**	-0.47	**I**	-0.35	**Se**	0.10	**B**
	Agg	0.60	**G**	0.17	**E**	-0.54	**Se**	0.06	**B**
	Rb	0.53	**G**	0.34	**E**	-0.45	**Se**	0.10	**B**
	Anx	0.63	**G**	-0.21	**I**	0.30	**Ot**	-0.39	**Ho**
	Sc	0.52	**G**	-0.19	**I**	0.11	**Ot**	-0.26	**Ho**
Parent	Wd	0.59	**G**	-0.11	**I**	0.38	**Ot**	-0.37	**Ho**
	Agg	0.68	**G**	0.32	**E**	0.21	**Ot**	-0.37	**Ho**
	Rb	0.65	**G**	0.46	**E**	0.12	**Ot**	-0.30	**Ho**
	Anx	0.46	**G**	-0.23	**I**	0.41	**Ot**	0.53	**Sc**
	Sc	0.43	**G**	-0.18	**I**	0.33	**Ot**	0.48	**Sc**
Teacher	Wd	0.31	**G**	-0.33	**I**	0.53	**Ot**	0.36	**Sc**
	Agg	0.35	**G**	0.66	**E**	0.05	**Ot**	0.38	**Sc**
	Rb	0.35	**G**	0.62	**E**	0.01	**Ot**	0.37	**Sc**
% Explained variance		**27%**		**14%**		**13%**		**10%**	

Note: *Anx* = Anxiety; *Sc* = Somatic complaints; *Wd* = Withdrawn behavior; *Agg* = Aggression; *Rb* = Rule-Breaking behavior, *G*=general; *Se*=self; *Ot*=other; *I*=internalizing; *E*=externalizing; *Ho*=Home; *Sc*=School; *B*=both.

Correspondence to easily computable scores

As shown in table 5, correlations between the algorithms for S, P, D and C and the component-scores derived from the PCA-model were all well above .90. This result was found with the T1-data as well as T2-data, indicating that these algorithms represent the component scores of the PCA model almost perfectly.

Table 5. *Correlations, based on T1-data and T2-data, between the PCA-components and the algorithms for Severity, Direction, Perspective and Context.*

PCA	Algorithm	T1	T2
Severity	mean (all subscores)	*0.99*	*0.99*
Direction	[mean (internalizing subscales) – mean (externalizing subscales)] / 2	*0.98*	*0.94*
Perspective	[mean (parent and teacher subscales) – mean (self subscales)] / 2	*0.96*	*0.93*
Context	[mean (teacher subscales) – mean (parent subscales)] / 2	*0.93*	*0.94*

Discussion

This study extended Kraemer and colleagues'(2003) multi-informant approach by modeling adolescents' internalizing and externalizing problems and their co-occurrence. We found support for a four component model of between-informant and between-domain convergences and divergences. The results support the hypotheses of this study: [1] between-informant discrepancies can be captured by a context (C) and a perspective (P) component; [2] between-domain discrepancies can be captured by a direction component (D); and [3] between-domain convergence can be captured by a severity component (S). This 'Multi-Informant Co-occurrence' model (MIC) was replicated with follow-up data from the same sample. The PCA-components almost perfectly correlate with the algorithms we developed for S, P, D and C. The formulas for these scores were directly derived from the hypotheses of this study. Therefore, this result can be interpreted as strong support for our main hypotheses. Provided that these results can be replicated in different samples, these algorithms allow for application of the model in research with small samples and individual diagnosis. While supporting the main hypotheses of this study, these results do not imply that the original model can be *replaced* by 4 Principal Components. We chose not to interpret more than 4 components, for reasons explained in the methods section. It was not expected that all variance could be explained by the MIC-model, as this is a model on broad-band specific features. To explain more of the variance the model would have to be extended to cover narrow-band specific features and more specific informant-discrepancies.

Using the MIC-model will increase both the validity and reliability of the measurement of emotional and behavioral problems. The reliability increases, because combining informants reduces measurement error. An extra advantage of aggregating information from multiple sources is that variance in the estimation of marginal parameters (e.g. prevalence, incidence) and associations (e.g. in regression analysis) is reduced (Hofler, 2005). Most importantly the validity increases by explicitly including the context (C) and perspective (P) in a multiple disorder model. Formulating and testing hypotheses regarding the reasons for discrepancies between specific informants is a more valid method for dealing with informant-discrepancies than only taking a mean score. Furthermore, the four-component MIC-model allows to take into account the high co-occurrence between the internalizing and externalizing domain by separating components with between-domain convergence (S, P and C) from a component with between-domain divergence (D).

The four components are not unidimensional scales of a specific trait, but should be interpreted in conjunction and as relative measures. The S-component can be used to discriminate between (pre)adolescents with regard to the amount of psychopathological problems. The D-component discriminates between those whose problems are relatively more internalizing and those whose problems are more externalizing. The P-component discriminates between those whose problems are relatively more self-reported, and those whose problems are more other-reported.

Finally, the C-component discriminates between those whose problems are relatively more observed at school, and those whose problems are more observed at home.

All four components contribute specific information regarding disorders and their co-occurrence. Although the relevance of each component may depend on the research question at hand, total disregard of some of these components implies using a reduced model. Although the relevance of each component may depend on the research question at hand, disregard of some of these components implies using a reduced model. Mean scores for INT and EXT, which are very often used in current research, can be transformed to S and D components without losing any information, using the provided algorithms. Using only INT and EXT mean scores , therefore implies a disregard of the P and C components.

The usefulness of the S- and D-components can be illustrated by the different ways in which temperament traits can influence the development of psychopathology (Shiner & Caspi, 2003). Such influences can be general, disorder-specific, or pathoplastic (Oldehinkel, Hartman, De Winter, Veenstra, & Ormel, 2004; Ormel, et al., 2005). The S- and D-component can be used to understand this issue of specificity better. General factors (e.g. frustration) are associated with the development of psychopathology, but not with the nature of problems. Thus, general factors should be positively correlated to the S-component, but uncorrelated to the D-component. Disorder-specific factors are associated with the development of a specific type of psychopathology, and therefore should not only be correlated to the S-component, but also to the D-component. Effortful control, for example, has been linked to the development of more externalizing problems in particular, and should therefore be positively correlated with the D-component, as a low D-score indicates relatively more externalizing problems. Pathoplastic factors are only associated with the nature, but not the amount, of problems, so these should only be correlated with the D-component. For example, shyness has been suggested as a component that is associated with relatively more internalizing problems, but not with the amount of problems. Accordingly, it should be positively correlated with the D-component, and uncorrelated with the S-component. An example of the application of the S- and D-components in the study of risk-factors can be found in Essex et al. (2006).

The P- and C-components contain additional information regarding the problems of (pre)adolescents. This information can, for example, be used to assess whether school-related factors (e.g. bad marks, truancy) are specifically related to problems at school, and therefore only associated with the C-component, or to more generalized problems, and therefore only associated with the S-component. Another example is the possibility to assess whether the P-component is related to the amount of psychosocial support a (pre)adolescent receives. A low score on the P-component indicates that problems are relatively more self-reported. These problems may remain untreated as the social environment does not recognize them as such. An example of the use of a context (C) and perspective (P) component is given by Perren et al. (2006) who show that differences in perspective between children and adults in hyperactivity/ impulsivity predict peer rejection.

The results provide rather strong support for the MIC-model. A strength of the study is, that it is based on a large cohort from the general population with limited non-response. A drawback is that vignettes were used to assess teacher ratings, instead of the original Teacher Rating Form (TRF; Achenbach, 1991b). The TRF has the advantage of using nearly the same items as the CBCL and YSR, so there might be less bias due to differences in the measurement-instrument used. However, the TCP-items follow the TRF-scales, and correlations between these measurements (Ferdinand, R.F., personal communication) indicate that the TCP-vignettes can be used as a proxy for the TRF scales. In the present study, we replicated the findings using follow-up data of the same sample, but replication in other samples is needed as well, among other things to assess the generalizability to other age groups and clinical populations.

The approach proposed in this article is preferable to a separate analysis of the subjective reports of different informants, because combining different perspectives and contexts results in a more objective measure of psychopathology. The approach is also preferable to separate analysis of the domains of internalizing and externalizing problems, because it allows to distinguish between common and broadband-specific effects. These are clear advantages for any research on psychopathology and especially for research on co-occurrence. The model could be used in clinical practice as all four components give important information regarding the condition of a patient which may be overlooked when looking to specific informant reports on specific disorders. It offers a clear method of combining different kinds of information that seems preferable to the situation where clinicians read all reports separately and combine this information subjectively. The algorithms we developed (see table 5) allow for easy, paper-and-pencil, computation of the S, P, D and C components. However, it is important to underscore that the interpretation of such scores at an individual level may not be straightforward (De Los Reyes & Kazdin, 2005), because some informant-discrepancies may have individual specific reasons.

As a final remark we would like to emphasize that the MIC-model is work in progress. The useful advantages and applications we have described are dependent on proper replications in other (e.g. clinical) samples. There are various possibilities for further development of the model. Firstly, the method is not specific to the instruments we used, nor to common mental health problems, nor to youngsters. Therefore, different measurement instruments can be used and the method can be expanded to other problem domains (e.g. ADHD symptoms, psychotic experiences, substance abuse). Secondly, using additional informants, for instance peers, would provide more information regarding different contexts or perspectives. Thirdly, future research may challenge some of the assumptions of the model. Context and perspective are defined as orthogonal components, but the way context and perspective shape an impression of someone and how that impression translates into questionnaire responses may not be fully captured by independent components, and deserves further investigation. Nevertheless, we expect this new tool to be useful for multiple purposes and encourage its application and further development.