Chapter 6

Stability and predictive utility of differences between informants

Abstract

Differences between informants challenge the quality of measurement and diagnosis, but might also provide useful diagnostic information. The current study was aimed at testing whether differences between informants showed rank-order stability during adolescence. Furthermore, it was tested whether these differences predicted changes in self-reported problems over time. The results showed large rank-order stability of the deviance of parent-reported problems from self-reported problems. Differences between informants were predictive of change in self-reported problems over time. These associations were probably too weak to be very useful for predictive purposes in clinical practice. It is discussed how the very stable deviances between self- and parent-reported problems that were found may be used in clinical practice for other purposes than prediction.

Introduction

A common finding in research on child- and adolescent psychopathology is that problem scores derived from different informants are only weakly correlated. This indicates that one or more of the instruments is unreliable, or that the instruments do not validly measure the same construct (Courvoisier, Nussbeck, Eid, & Cole, 2008) or a combination of these. If instruments are unreliable then each measurement can be regarded as an imprecise piece of information about the same construct. If they are not valid measures of the same construct, then each measurement may be regarded as a piece of information about a different construct. Related to this second possibility, it has been argued that information from each informant can be regarded as adding specific diagnostic information (Achenbach, et al., 1987). Even discrepancies between different informants as such may contribute to a more complete picture of an individual's functioning. To test the relevance of this perspective several questions need to be addressed, two of which are the focus of the current article. The first question is how discrepancies develop over time. Do they show stability or do they vary across measurement occasions? The second question is how useful discrepancies are in terms of predictive value for the development of self-reported behavioral and emotional problems.

Specific hypotheses with regard to the above-mentioned questions can only be developed if we have an idea of what is actually measured by discrepancies. The most general and obvious feature of informant discrepancies is that they demonstrate that

multiple answers can be given to the same questions about the same individual. We distinguish four specific aspects that are likely to influence this variability. First, differences between informants can indicate uncertainty about the answer to a question, so that informants may make different guesses. Second, differences can indicate that the actual behavior to which the questions refer is variable and only expressed in certain contexts or at certain moments (e.g. Kraemer, et al., 2003). Third, differences can indicate that some behaviors, for instance emotions that are not clearly expressed, are difficult to observe for some informants (e.g. Harkness, Tellegen, & Waller, 1995). Fourth, differences can indicate that informants have a different judgment about the same observations (see De Los Reyes & Kazdin, 2005). In sum, we assume that differences between informants indicate informant uncertainty, variability of behavior, low observability, or differences in judgment. In our view, these four general terms cover most of the more specific causal pathways that have been described in the literature. For example, differences between self- and parent-report may be related to denial of problems (Harkness, et al., 1995) or oppositionality, which can be regarded as differences in judgment or visibility. Also, it has been suggested that differences may be related to differences in the context in which a subject is observed (De Los Reyes & Kazdin, 2005), which can be regarded as involving variability of behavior over contexts. Finally, the four terms also include idiosyncratic reasons for differences between informants, which can be viewed as informant-specific interactions between observations, judgments and responses.

Variability and low observability are behavior properties that may change during adolescence. For example, observability may increase due to communication between parents and adolescents. It may also decrease due to the fact that an adolescent spends less time at home. However, at the same time we can expect stability: if a behavior-pattern is very context-dependent or difficult to observe for others at one measurement occasion, it is likely that the same behavior-pattern will also be more variable over context or difficult to observe at a later measurement occasion. Furthermore, the context and perspective of a specific informant and the manner in which informants judge about behaviors, can be expected to have some stability as well. For these reasons, informant discrepancies are expected to show a substantial amount of stability. This hypothesis is supported by previous reports on the stability of informant-specific factors and differences between informants in childhood (e.g. Courvoisier, et al., 2008; Grimm, Pianta, & Konold, 2009).

With regard to the question how useful informant discrepancies are as a predictor of the course of self-reported problems, we expect larger differences to be associated with more variability over time. If a response is given with uncertainty, the question is apparently not very clear, which increases the likelihood that the respondent will give a different response at a future occasion. If the behavior is variable, it is relatively likely to change over time due to changes in context like going to another school. This is especially true in adolescence, where such changes are common. If there are

differences in judgment, there is some chance that the self-report will change because of the influence of the judgment of others.

With respect to the direction of changes over time, we hypothesize that in case of informant discrepancies, self-reported ratings will change in the direction of other-reported ratings. More specifically, if self-reported scores are low compared to other-reports, we expect them to increase, and if they are relatively high, we expect them to decrease over time. If discrepancy indicates a range of possible responses, it is more likely that a response will change into an alternative that has been given previously by another informant than that a completely new response will be given. At the same time, we expect that more discrepancies, regardless of whether the other informant reports more or less, predict a decrease of self-reported problems over time. Thus, we expect *non-directional* discrepancies to predict a decrease of self-reported problems and we predict *directional* discrepancies to predict a change towards the amount of problems reported by the other informant. The rationale for this is that we assume that discrepancies indicate uncertainty about and variability of the problems that the items ask about. We expect that these problems are more likely to involve temporary, context-dependent, issues rather than stable traits and are therefore more likely to disappear at later ages.

The predictive value of informant discrepancies has been suggested by results of Pelton and colleagues (Pelton & Forehand, 2001; Pelton, Steele, Chance, Forehand, & Family Hlth Project Res, 2001), who reported that differences between self- and parent-reports on the quality of the parent-child interactions predicted internalizing and externalizing problems at a later age. Furthermore, differences between self- and parent-reported symptoms have been found to predict negative outcomes like problems at school, police contacts and drug abuse (Ferdinand, van der Ende, & Verhulst, 2004, 2006). Others have reported that differences between self- and parent-reported parenting style were predictive of depressive symptoms and anti-social behavior (Feinberg, Howe, Reiss, & Hetherington, 2000), and of internalizing problems and social competence (Guion, Mrug, & Windle, 2009).

In sum, we aimed to investigate several expectations with respect to the stability and predictive utility of differences between informants. These expectations were tested on data from a longitudinal study of (pre)adolescents. We used self- and parent-reports of behaviors and emotions related to aggression on the one hand and withdrawal from social contact and depression on the other. For both domains we expected to find stable differences between informants and utility of these differences in predicting variability in self-report over time.

Methods

Sample

Subjects were participants in the 'Tracking Adolescents' Individual Lives Survey' (TRAILS), a prospective multi-cohort study of Dutch (pre)adolescents. The study

involved a representative sample from the general population and is described in detail in Huisman et al. (2008). Briefly, the 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, mean age = 11.09, SD =0.55) 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). To date, the population cohort has been assessed three times (T1: March 2000- July 2001, T2: September 2003- December 2004, T3: September 2005-December 2007). Participation rates were 96.4% at T2 (mean age= 13.55, SD = 0.53), and 81.4% at T3 (mean age= 16.25, SD = 0.73). After complete description of the study to the subjects, written informed consent was obtained from the parents at each assessment wave and from the adolescents at T2 and T3. T1, T2, and T3 data are used in the present study.

Instruments

The Dutch versions of the Youth Self-Report (YSR; Achenbach, 1991c; Verhulst, et al., 1997b) and the Child Behavior Checklist (CBCL; Achenbach, 1991a; Verhulst, et al., 1996) were used to assess self- and parent-reported behavioral and emotional problems. The CBCL and YSR are 112-item questionnaires on which emotions and behaviors are rated on a 3-point rating scale (not [0], sometimes [1], or very often [2]). The period over which respondents are asked to report is the last six months. In the TRAILS-study the questionnaire was completed by one of the parents, which was the mother in most cases. Factor analysis on these items revealed a structure of eight syndrome scales (Achenbach, 1991c). In the current study the scales Withdrawn-Depressed (Wd; 8 items, CBCL: α=0.71, YSR: α=0.64) and Aggressive Behavior (Agg; 18 items, CBCL: α=0.88, YSR: α=0.82) will be used.

Statistics

All analyses were done using MPlus version 5.2. Because Wd and Agg are not normally distributed we used maximum likelihood estimation with robust standard errors (MLR), which is relative robust to deviations from normality. We used the Root Mean Square Error of Approximation (RMSEA) and the Comparative Fit Index (CFI) to evaluate model fit. An RMSEA below .05 and a CFI above .95 were regarded as indicating adequate fit.

• *Defining differences between informants*

We used two distinct measures of the differences between informants. On the one hand we developed a measure of the absolute amount of discrepancies that are observed between informants (DISi). The DISi scores were based on a summation of all discrepancies between responses on all items of a subscale:

$$DISi = \Sigma_{i=1\ldots s} (|\text{Item i self} - \text{Item i parent}|).$$

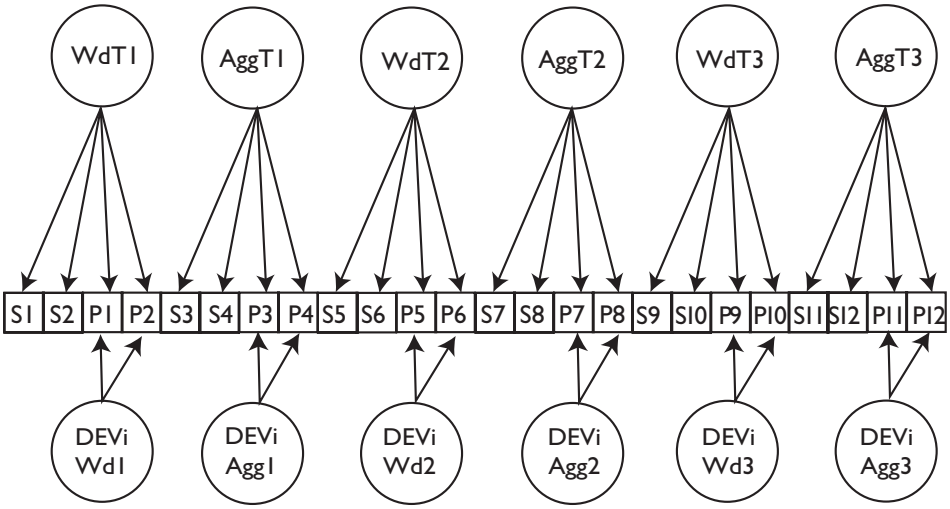

Figure 1. The CT-C(M-1) model for three occassions of measurement of subscales Wd and Agg.

Note: Correlations between factors are excluded for simplicity. All factors are correlated except for self-report and parental report at each measurement wave (T1-T3). *Wd*=Withdrawn-Depressed; *Agg* = Aggressive Behavior; *DEVi* = deviance of other informant (parent); *S1-S12* indicators of self-reported problems; *P1-P12* indicators of parent-reported problems.

On the other hand we used a measure of the 'deviance' of the report of another informant from self-reported problems (DEVi) on the basis of a recently developed latent model for multi-informant data. This measure is high when the other informant reports more problems, and low when the other reports less problems. The DEVi scores were derived from the "Correlated Traits – Correlated Methods minus one model" (CT-C(M-1); Eid, et al., 2003). Overviews of this model and arguments for why it is a good approach for modeling multi-informant data are given elsewhere (Courvoisier, et al., 2008; Eid, et al., 2008). For the current article this model (see figure 1) was used to estimate the unique variance of parental reports, i.e. the information that is not predictable from self-report. To develop a CT-C(M-1) model one has to choose a reference-method, which in this case was self-report. Subsequently, one estimates 'Trait' factors on which indicators of all informants have

loadings, and 'Method' factors on which indicators of only one informant have loadings. These 'Method' factors are estimated for all informants except for the reference-method and are uncorrelated to the 'Trait' factors. Therefore, these factors are best interpreted as unique deviance of other informants from the reference-method (Geiser, Eid, & Nussbeck, 2008). In the current article we chose self-report as a reference-method and used the CT-C(M-1) model to estimate deviance scores of parental report (DEVi). Unfortunately, fitting the model on the basis of items as indicators resulted in a too complicated and non-converging model. As an alternative we used a split half method (see Courvoisier, et al., 2008). We computed indicators on the basis of half of the items of each subscale, which resulted in two indicators for each subscales at each assessment wave. These indicators were used to fit the CT-C(M-1) model.

Our DISi measure reflects all observed discrepancies on a subscale regardless of the direction of these discrepancies. Earlier approaches to define non-directional discrepancy have used the absolute difference of scale scores rather than items (Feinberg, et al., 2000; Pelton & Forehand, 2001). This approach is likely to correlate with ours, but in our view the item-level approach is more consistent in capturing all non-directional discrepancies between informants. Low scores indicate few discrepancies, high scores indicate many discrepancies. It appears that the DISi scores are positively correlated to the variables estimated in the CT-C(M-1) model. This can be understood as indicating that if more problems are observed the amount of discrepancy will generally be higher. However, variance in DISi cannot be completely explained by the CT-C(M-1) model. As the CT-C(M-1) approach provides a concise latent variable model for multi-informant data we were only interested in the variance in DISi that was not explained by the CT-C(M-1) model. Therefore, in all analyses of DISi we controlled for the influence of the CT-C(M-1) factors.

• *Defining variability over time*

We used two different variables to capture differences in self-reported problems between multiple occasions of measurement (T1, T2 and T3). First, we used a linear growth model (see figure 2). In a linear growth model an individual growth curve is estimated for each individual, which is characterized by two variables: an intercept and a slope. We specified the slope as the linear deviance from the T1-measure, which will be referred to as deviance over time (DEVt). High values indicate an increase of self-reported problems over time, while low values indicate a decrease.

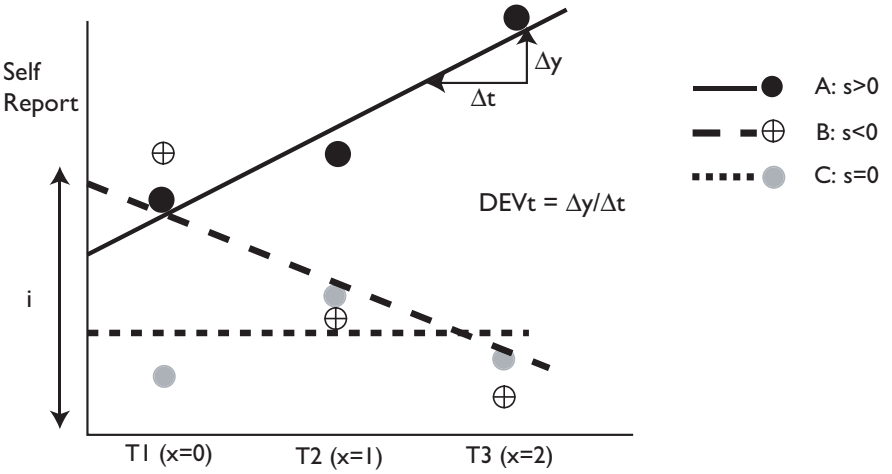

Figure 2. Examples of the individual growth-curves of three individuals (A-C).
Note: i = intercept; *DEVt* = deviance over time; *T1-T3* = measurement waves; *A* = individual with increase of self-reported problems over time; *B* = individual with decrease; *C*=individual with no linear change.

The DEVt variable can be used to investigate whether differences between informants are predictive of increases or decreases of self-reported problems over time. However, we were also interested in whether differences between informants were predictive of any variability over time, regardless of the direction of this variation. For this reason we developed a variable indicating the amount of 'discrepancy over time' in item-responses between T1 and the other occasions of measurement:

$$DISt = \Sigma_{i=1\ldots s} (|Item\ i\ T1 - Item\ i\ T2| + |Item\ i\ T1 - Item\ i\ T3|).$$

High values indicate many discrepancies between self-reported problems over time, while low values indicate few discrepancies.

• *Stability and predictive utility*

We investigated the rank-order stability of the variables DEVi and DISi over the three occasions of measurement by computing test-retest correlations between different occasions of measurement (T1, T2, T3). Subsequently, we used linear regression analysis to test whether DEVi and DISi predicted DEVt and DISt. First, we expected DISi to be a predictor of DISt. That is: we expected that more discrepancies between informants at T1 would predict more discrepancies between T1 self-reported

problems and self-reported problems at T2 and T3. Second, we expected DEVi to predict DEVt, because we expected that most change would occur in the direction of the problems reported by the other informant. That is: we expected that if parents reported more problems at T1 then self-report would increase over time. Third, we expected DISi to predict DEVt, because we expected discrepancies to be an indicator of instability, unreliability or context-dependency of the reported behaviors. That is: we predicted that more discrepancies at T1 would predict a decrease of self-reported problems over time.

• *Preventing chance findings*

Because we fitted multiple latent variable models and tested several linear regression models we felt that there was a risk of chance capitalization. For this reason we randomly split the original sample in two parts. In the first sample (N=1099) we developed the latent variable and regression models and we replicated these models in the second sample (N=1124).

Results

The CT-C(M-1) model

The CT-C(M-1) model with self-report as reference method fitted well to the data (χ^2 =631.5; *df* = 183; RMSEA<.05; CFI > .95). Factor-loadings of this model are reported in table 1. For each measurement-wave (T1, T2, T3) and each subscale this model resulted in a distinction between reference factors, which are strongly influenced by self-report, and unique deviance factors (DEVi) which are only influenced by parental report and uncorrelated to the reference factors. The DEVi factors were calculated for both the Aggressive Behavior and the Withdrawn-Depressed subscales and were used in subsequent analyses.

Growth model of self-reported problems

A growth model was fitted that included T1-T3 self-reports of the subscales Withdrawn-Depressed and Aggressive Behavior. Factor-loadings are reported in table 2. This model showed adequate fit-indices (χ^2 =16.8; *df* = 6; RMSEA<.05; CFI>.95). The model resulted in estimated linear growth curves for each individual for both the Aggressive Behavior (Agg) and Withdrawn-Depressed (Wd) scales. These are indicated by individual-specific intercept-variables and slope-variables. The slope-variables capture individual differences in linear change during adolescence (DEVt).

Table 1. *Standardized loadings of the CT-C(M-1) model for Withdrawn-Depressed and Aggressive Behavior:*

Informant	indicator	T1 Wd	T2 Wd	T3 Wd	DEVi T1 Wd	DEVi T2 Wd	DEVi T3 Wd
Self	Wd1	0.77	0.73	0.81			
	Wd2	0.67	0.72	0.79			
Parent	Wd1	0.21	0.33	0.31	0.75	0.76	0.78
	Wd2	0.24	0.35	0.35	0.67	0.68	0.72
		T1 Agg	T2 Agg	T3 Agg	DEVi T1 Agg	DEVi T2 Agg	DEVi T3 Agg
Self	Agg1	0.86	0.84	0.84			
	Agg2	0.84	0.81	0.88			
Parent	Agg1	0.35	0.38	0.37	0.84	0.82	0.81
	Agg2	0.34	0.40	0.40	0.84	0.83	0.84

Note: The indicators are derived from splitting the subscales in halves for each informant at each assessment wave. *Wd* = Withdrawn-Depressed; *Agg* = Aggressive Behavior; *DEVi* = Deviance of parental report from self-report.

Table 2. *Standardized loadings of the growth-model for self-reported Withdrawn-Depressed and Aggressive Behavior.*

	Intercept	DEVt Wd
Wd T1	0.687	
Wd T2	0.682	0.293
Wd T3	0.637	0.548
		DEVt Agg
Agg T1	0.674	
Agg T2	0.699	0.269
Agg T3	0.678	0.523

Note: The indicators are derived from splitting the subscales in halves for each informant at each assessment wave. *Wd* = Withdrawn-Depressed; *Agg* = Aggressive Behavior; *DEVt* = linear change relative to T1.

Rank-order stability of differences between informants

The rank-order stability of differences between informants was investigated by computing test-retest correlations of the DEVi and DISi variables for both the Aggressive Behavior and Withdrawn-Depressed subscales. As reported in table 3, all test-retest correlations for the DEVi variables were above .75. This indicates that the deviance of parental report from self-report tends to be very stable during

adolescence. DISi on the other hand was much less stable, with test-retest correlations between .09 and .25. These are partial correlations, because at each measurement-wave (T1, T2, T3) we controlled for self-reported problems and DEVi. Thus, the deviance (i.e. higher or lower estimation) of parental reports from self-report was found to be very stable over time, but the amount of discrepancy in item responses much less so.

Table 3. *Test-retest correlations of differences between informants.*

Subscale	Informant difference	T1 - T2	T2 – T3	T1-T3
Wd	DISi	.15	.19	.09
	DEVi	.82	.87	.80
Agg	DISi	.25	.24	.19
	DEVi	.80	.79	.77

Note: *Wd* = Withdrawn-Depressed; *Agg* = Aggressive Behavior; *DISi* = Discrepancies between informants; *DEVi* = Deviance of parental report from self-report.

Table 4. *Differences between informants as predictors of variability in self-reported problems over time.*

Subscale	Variability	T1 predictors	B	SE	*Beta*	R^2
Wd	DISt	Self-report	5.85	0.41	*0.53**	.43[a]
		DISi	0.32	0.07	*0.20**	
		DEVi	-0.02	0.12	*0.00*	
						.46
	DEVt	Self-report	-0.17	0.06	*-0.14**	.05
		DISi	-0.02	0.01	*-0.14**	
		DEVi	0.09	0.02	*0.20**	
						.09
Agg	DISt	Self-report	12.66	0.81	*0.56**	.48
		DISi	0.41	0.06	*0.26**	
		DEVi	-0.23	0.10	*-0.07*	
						.51
	DEVt	Self-report	-0.75	0.12	*-0.24**	.09
		DISi	0.26	0.04	*-0.09**	
		DEVi	0.10	0.02	*0.22**	
						.12

Note: *Wd* = Withdrawn-Depressed; *Agg* = Aggressive Behavior; *DISt* = discrepancies in self-report between T1 and other occasions of measurement (T2, T3), *DEVt* = deviance (linear increase/decrease) of self-reported problems over time, *DISi* = discrepancies between informants; *DEVi* = Deviance of parental report from self-report.

[a] The R^2 after self-report shows the amount of explained variance if only self-report was used as a predictor. It does not correspond to the regression-coefficients, which are all derived from the analyses with multiple predictors.

* $p<.05$

Predictive utility of differences between informants

Linear regression analysis was used to test whether differences between informants (DEVi and DISi) were predictive of variability over time (DEVt and DISt). The results are shown in table 4 and were very similar for the Agg and Wd subscales. First, as expected DISi was a positive predictor of DISt. This indicates that more discrepancies in item-responses between informants at T1 predict more discrepancies of self-reported problems between T1 and later occasions of measurement (T2 and T3). Second, DEVi was a positive predictor of DEVt. This indicates that the directional deviance of parental reports from self-report was predictive of the direction of change in self-reported problems over time. Thus, as expected, higher parental reports predict an increase of self-reported problems and lower parental reports predict a decrease. Finally, DISi was a negative predictor of DEVt. This indicates that, confirming our hypothesis, more discrepancies in item responses are predictive of a decrease in self-reported problems over time.

Thus, both DEVi and DISi predicted variability in self-reported problems over time. However, the amount of explained variance was small. As can be observed from table 2, differences between informants (DEVi and DISi) did not substantially increase the amount of explained variance in variability over time (DEVt and DISt).

Replication

The above presented results were all based on analyses of a random half of the total sample. All analyses were replicated with the other half. The results replicated very well. However, one discrepancy was found (all other results can be obtained from the authors upon request). The growth-model of self-reported problems showed slightly worse model-fit in the replication sample (χ^2 =16.8; *df* = 6; RMSEA=.061; CFI>0.95). This shows that the linear growth-model did not completely capture the development of problems during adolescence, but does not invalidate the results presented in this paper because the interpretation of the slopes (DEVt) is not altered by it.

Discussion

It has been suggested that discrepancies between informants may contain useful diagnostic information (e.g. De Los Reyes, & Kazdin, 2005). The aim of the current study was to investigate the stability and predictive utility of this information. Strong support was found for rank-order stability of the deviance between parental report and self-report. Furthermore, informant discrepancies predicted variability in self-reported problems during adolescence. However, predictive utility was small: differences between informants accounted for little explained variance in variability of self-reported problems during adolescence. In the following these results will be used to evaluate whether informant differences are interesting psychological phenomena that contain useful information, and whether this information can be used in clinical practice. Before discussing these three issues we will mention some strengths and limitations of the study.

Strengths and limitations

A particular strength is that differences between informants were captured by two measures. In our view these two measures are well-suited to cover the full domain of informant discrepancies. On the one hand the CT-C(M-1) model adequately captures systematic deviances of other informants from self-report. On the other hand the summarized discrepancies between item-responses capture non-systematic non-directional discrepancies. Furthermore, we used two different measures of variability over time. We think that these two measure together adequately capture both the systematic change and non-systematic variability of self-reported problems during adolescence. Altogether, we think that these four variables allow for a rather complete understanding of the associations between informant differences and variability over time.

The generalizability of the study is obviously limited by the fact that it only covers self-reports and parental reports during adolescence in one specific longitudinal study (TRAILS) using one specific measurement approach (CBCL and YSR). Furthermore, we only evaluated the utility in predicting variability of self-reported problems over time. This outcome was useful in evaluating the diagnostic value of differences between informants, but a complete evaluation of predictive utility should capture other outcome variables as well.

Can differences between informants be regarded interesting psychological phenomena?

This study clearly showed stability of the deviance between self- and parent-reported problems during adolescence. Obviously, these impressively stable deviances deserve scientific attention. They can be completely interpreted as involving stability of the reports of each single informant. That is, each informant-report is by far the best predictor of the report of that same informant at a later occasion. This does certainly not imply that these reports are only influenced by informant-characteristics or 'bias' (van der Valk, van den Oord, Verhulst, & Boomsma, 2001).

Furthermore, it was found that differences between informants were predictive of variability in self-reported problems over time. Therefore, differences between informants at T1 may be interpreted as contributing diagnostic information about self-reported problems at T1. More discrepancies predict that the T1 self-report is somewhat more likely to change. Agreement between multiple informants was a predictor of stability, disagreement was a predictor of change.

What kind of diagnostic information is captured by differences between informants?

The study suggests that differences between informants primarily reflect the stability of each unique informant's report. This uniqueness is probably related to the specific perspective and context of that informant (De Los Reyes, 2009; Kraemer, et al., 2003; Noordhof, et al., 2008). However, the selves and parents involved are also

unique individuals with a unique way of accomplishing the task of completing a questionnaire. There are many ways to do this and the informants do not get feedback on how well they performed. All non-random idiosyncratic ways of observing and responding may contribute to a stable deviance between two unique informants. Therefore, the most important diagnostic information that is contained in differences between informants is that the instruments do not reliably measure exactly the same underlying quantitative construct. Assuming that none of the informant-reports can be accepted as a gold standard measure this implies that the instruments do not permit a very precise 'unbiased' estimation of the actual amount of psychological problems. This does not mean that the two instruments (self-report and parent-report) may not give very precise estimations of two different quantitative constructs related to two different perspectives (self vs parent).

Differences between informants were somewhat predictive of systematic change and non-systematic variability of self-reported problems over time. The associations found were in the direction expected on the basis of our hypotheses. Thus, the results do not contradict the very broad interpretations of differences between informants given in the introduction (uncertainty, variability, observability, judgment). However, the small associations found and the large amount of possible explanations for these findings do not permit strong conclusions regarding what is measured by differences between informants.

Can differences between informants be used in clinical practice?

We did not find evidence that discrepancies between informants have much utility for predicting the variability in self-reported problems during adolescence, because the amount of explained variance was too small. As discussed by McGrath (2008) in situations of practical decision-making a 'best single predictor' heuristic may often be preferable to taking into account multiple predictors of an outcome, which is specifically the case with weak predictors that add only little explained variance.

However, we do not think that the utility of differences between informants should only be evaluated in terms of predictive utility. Specifically, we think that differences may have utility both in diagnosis and in therapeutic interventions. The multiple instruments used in clinical practice do not result in a precise estimation of the actual amount of problems, but in multiple rather stable pieces of information, which may be imagined as points on a hypothetical range. This range, which will show much stability over time, may currently be a more realistic conceptualization of psychopathology than the quest for point-estimates. Accepting this viewpoint in clinical practice means that an estimate given by one informant is regarded as only one point on a hypothetical range. Adding more instruments, repeated measures and more informants does not help to reduce this range. On the contrary, it helps to further characterize the range of estimates that can be given for the same person. Communicating about ranges and uncertainties can increase the reliability of diagnostic information and using differences between informants can be one way to accomplish this goal. A very interesting and

comparable idea has been proposed for conceptualizing the range of possible change (RPC) supported by clinical trials (De Los Reyes & Kazdin, 2006).

Differences between informants may also have therapeutic utility. Using self-report as a reference-measure a clinician may discuss deviances of other reports from self-reported problems and aim to understand the idiosyncratic reasons for these differences. While some of these may just involve differences in understanding items or response-styles, other deviances may result from a real difference between contexts (e.g. home vs school), between self-judgment and judgment by others or a lack of insight by some informants. Such post-hoc individualized explanations will currently not result in explicit and quantifiable knowledge (see McGrath, 2008), but may increase diagnostic understanding and therapeutic effectiveness. Thus, differences between self-report and other informants might be used as a therapeutic tool. Such interpretations can be further advanced by studies that aim at understanding the multiple causes underlying discrepancies between informants (e.g. De Los Reyes, 2009).

Conclusion

Stable differences between informants reflect the stability of unique informant reports, which are well captured by the CT-C(M-1) model. More discrepancies between informants are predictive of more variability in self-reported problems over time, which is mostly in the direction of the parental report. In this study predictive utility was not impressive, but a full evaluation of predictive utility should be based on more outcomes and on comparisons of multiple predictive algorithms. Furthermore, differences between informants can have important diagnostic and therapeutic utility.