



## SUMMARY AND CONCLUSIONS

### **Preliminary Study Convergent Validity LIJ**

*Authors: dr. I.J.L. Egberink and prof. dr. R.R. Meijer*

#### **Answers to the Research Questions**

The National Set of Instruments Juvenile Justice System (in Dutch: *Landelijk Instrumentarium Jeugdstrafrechtketen*; hereafter referred to as LIJ) is a set of tools designed for young people, who at the age of 12 through 17 have been in trouble with the police concerning a criminal offense, and this set of tools can be used by different chain partners. The LIJ maps both the probability of a repeat offense and issues for concern in order to find a suitable (care) intervention for the young person. The LIJ has a scientific basis and the aim is an assessment of the LIJ by the Dutch Committee on Tests and Testing Affairs (in Dutch: *Commissie Testaangelegenheden Nederland*, hereafter referred to as COTAN). Therefore, it is important that already (partly) conducted research regarding the psychometric quality of the ‘test-parts’ of the LIJ will be completed on the basis of the COTAN-criteria. Regioplan (Timmermans & Witvliet, 2011) tried to investigate the convergent validity of the care signals of the LIJ by comparing the LIJ with the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001; van Widenfelt, Goedhart, Treffers, & Goodman, 2003). However, at the same time Boonekamp (2010) concluded in his research regarding the SDQ that there are doubts about the reliability and validity of the SDQ for use with youth offenders. As a result, in the present study we evaluated past research with respect to the convergent validity and future possibilities to determine the convergent validity of the LIJ.

The central question of this preliminary study is: “Is it possible to investigate the convergent validity of the LIJ?”. To answer this question, several available documents and research reports with respect to the development of the LIJ and earlier research regarding the LIJ were studied. Based on this information, additional documents were collected and studied. Furthermore, all items of the different parts of the LIJ were inspected and evaluated

to investigate whether a combination of some items or a combination of some parts of the LIJ could be considered as one scale. The LIJ was also evaluated at a general level to investigate what construct or which constructs the LIJ pretends to measure. This information was used to conduct a literature study into other questionnaires that pretend to measure the same as the LIJ. Furthermore, the current state of affairs in the literature concerning the concept of validity was examined, in particular the concept of convergent validity, with a special eye towards the evaluation of the LIJ according to the COTAN-criteria. After considering all the available information, it seems that the central research question cannot be answered with a simple ‘yes’ or ‘no’.

The answer to the first sub-question “Which parts of the LIJ can be considered as one scale?” is that the LIJ, specifically the Dynamic Risk Profile (DRP), does not contain parts that could be considered as one scale. However, (detecting) psychological dysfunctioning as part of ‘Care Assessment’ can be considered as one scale, under certain conditions (for a differentiation, see section ‘Recommendations for validity research with regard to ‘Care Assessment’ and the SDQ and/or SPsy’). The LIJ as a whole, in particular the DRP, is a risk assessment instrument.

The answer to the related sub-question “Which questionnaire converges with the LIJ?” is more complicated. Assessing the (convergent) validity is only relevant for the LIJ as a whole and the part ‘Care Assessment’. This will be discussed in the following sections ‘Recommendations for validity research with regard to the LIJ as a whole’ and ‘Recommendations for validity research with regard to ‘Care Assessment’ and the SDQ and/or SPsy’.

### **Recommendations for Validity Research with Regard to the LIJ as a Whole**

Although, there are different risk assessment instruments available in the Netherlands, yet little validity research (and sometimes also research regarding norms) is conducted with these kinds of instruments and/or the instruments are not intended for youth offenders aged 12 through 17. Therefore, research into the convergent validity of the LIJ seems impossible at this moment. Given the current view on validity, it might be more useful to collect validity information with regard to the purpose of the instrument, namely discriminating between low, medium, and high risk at recidivism (screening), applying suitable care interventions to the young people (diagnostics), and evaluating the effectiveness of those intervention assignments (evaluation). Even though that kind of research is outside the scope of this

preliminary study, it is important and therefore some suggestions will be made for such research.

The Recidivism Monitor of the Research and Documentation Center (in Dutch, *Wetenschappelijk Onderzoek- en Documentatiecentrum*) can play an important role in this type of validity research, especially for screening and evaluation purposes. However, because only the cases that are handled by the Public Prosecutor are kept in this monitor, one should be cautious for the so-called ‘restriction of range’ problem. Therefore, it is important to have as much spread as possible in the seriousness and frequency of the offenses. The phased introduction of the LIJ could be used to investigate its effectiveness. For example, how much and how often young people assessed by the LIJ re-offended in comparison to young people not assessed by the LIJ. Furthermore, research from van der Knaap and Alberda (2009) regarding the predictive validity of the Recidivism Risk Assessment Scales (in Dutch, *Recidive Inschattingsschalen*) can be used as a guideline.

### **Recommendations for Validity Research with Regard to ‘Care Assessment’ and the SDQ and/or SPsy**

The answers to the items with regard to the psychosocial situation of the young person from the ten different domains together form the part ‘Care Assessment’. Based on these answers, the user is asked to provide a clinical judgment regarding the extent of concern for each domain, regarding the general psychosocial situation, and regarding the risk of harming himself/herself. Also the SDQ and/or Screening instrument psychological disorders (in Dutch, *Screeningsinstrument Psychische stoornissen*, hereafter referred to as SPsy) are filled out by the young person and/or his/her parent(s) to detect psychological dysfunctioning.

Standardization is an important part of data collection. Standardization is sometimes difficult with clinical judgments. Therefore, standardized instruments without clinical judgments are preferable to standardized instruments with clinical judgments (Bosker, 2008). Furthermore, with clinical judgments the interrater and intrarater reliability should be investigated (Regioplan investigated the interrater and intrarater reliability, see section 4.1 Consideration at the Item Level of the report).

Because on top of the items with regard to care signals also the SDQ and/or SPsy are filled out to detect psychological dysfunctioning and because there is an overlap between those parts, an alternative is to replace the clinical judgments with one questionnaire that maps the psychosocial problems. After all, the aims of administering the SDQ and/or SPsy and the items from the LIJ for the purpose of ‘Care Assessment’ are similar. On the one hand, the

validity of both the SDQ and the SPsy can be investigated more extensively to deal with the content overlap with 'Care Assessment'. On the other hand, another solution may be the use of another already (partly) validated questionnaire. Several questionnaires are available to detect psychological dysfunctioning and the reliability and validity of those instruments have been assessed by the COTAN (e.g., the Dutch questionnaire "Sociaal-Emotionele Vragenlijst"). However, most of these questionnaires are not validated for youth offenders. A suggestion is to choose a questionnaire that is evaluated positively by the COTAN that is similar to the SDQ and/or SPsy. The research for the original population can then be replicated for the population of youth offenders.