The Assessment of Executive Functioning Using the Childhood Executive Functioning Inventory (CHEXI)

Lisa B. Thorell and Corinne Catale

Development of the CHEXI

The strong association between deficits in executive functioning (EF) and psychiatric disorders such as Attention-Deficit Hyperactivity Disorder (ADHD) has led to the development of a myriad of neuropsychological tests designed to capture both executive functioning in general and specific abilities within the EF domain. However, relatively few EF rating instruments are available. Compared to neuropsychological tests, rating instruments generally capture more global aspects of behavior. Although ratings have the disadvantage of suffering from rater biases, they have the advantage of capturing behavior over an extended period of time and in different settings (e.g., home, school). As they are easy to administer, ratings can also be most valuable as a screening instrument for identifying children at risk of developing psychiatric disorders or functional impairments such as poor school achievement. In addition, EF rating instruments may be used to identify children with EF deficits, with or without psychiatric disorders, who may be helped through targeted intervention programs focused on training executive functions (e.g., Klingberg et al., 2005; Thorell, Lindqvist, Bergman Nutley, Bohlín, & Klingberg, 2009).

The most commonly used EF rating instrument for children is the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia, Andrews Espy, & Isquith, 2003; Gioia, Isquith, Guy, & Kenworthy, 2000; see also Chap. 18 in this volume), and more recently the Barkley Deficits in Executive Functioning Scales—Child and Adolescent Version (BDEFS-CA; Barkley, 2012; see also Chap. 15 in this volume) and the Comprehensive Executive Functioning Inventory (CEFI; Naglieri & Goldstein, 2013; see also Chap. 14 in this volume) have been introduced. The BRIEF has been found to be very useful in distinguishing between, for example, children with ADHD and normally developing controls (e.g., Gioia, Isquith, Kenworthy, & Barton, 2002). However, one potential limitation of the BRIEF is that in addition to measuring executive functioning, this scale also directly measures ADHD symptoms in that items such as “is impulsive” and “has a short attention span” are included, and these are almost identical to the symptom criteria for ADHD (APA, 1994). Scales like this could of course be of great use in both research and clinical practice, but the semantic overlap between the items included in the BRIEF and the symptom list for ADHD means that erroneous conclusions about EF deficits in children can be drawn. As repeatedly shown (e.g., Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005), not all children with ADHD have EF deficits and vice versa. It may therefore be valuable to be able to

L.B. Thorell (✉)
Karolinska Institutet, Stockholm, Sweden
e-mail: lisa.thorell@ki.se

C. Catale
University of Liège, Liège, Belgium

S. Goldstein and J.A. Naglieri (eds.), Handbook of Executive Functioning.
distinguish between EF deficits and ADHD symptoms. This may be especially important in relation to functional impairments, as different combinations of ADHD and/or EF deficits may be differentially related to, for example, poor academic achievement (Biederman et al., 2004).

Another potential problem is that the BRIEF (63–86 items depending on the version used), the BDEFS-CA (70 items), and the CEFI (90 items) are all long questionnaires. A short form (20 items) of the BDEFS-CA is available, but no factor analysis of this version is presented in the manual, indicating that the short version is best used as a general EF scale rather than as a measure with subscales tapping specific EF functions. Finally, some items included in the BDEFS-CA would seem to be more suitable for older compared to younger children. This is especially true for the time management subscale, which includes statements such as “wastes or doesn’t manage his/her time well” or “has difficulty judging how much time it will take to do something or to get somewhere.”

In order to address the limitations of available questionnaires intended to tap into EF, the Childhood Executive Functioning Inventory (CHEXI; Thorell & Nyberg, 2008) was developed as a quick screening instrument specifically targeting different types of executive control rather than including more general statements or items directly related to the symptom criteria for ADHD. In the present chapter, we provide an overview of previous studies using the CHEXI in several different countries in Europe, Asia, and South America, and we also present some new, unpublished data.

### Description of the CHEXI

The CHEXI first included 26 items, but in the original study introducing the instrument (Thorell & Nyberg, 2008), two items were excluded due to too low sampling adequacy. The most commonly used version therefore has 24 statements and takes about 5–10 min to complete. The complete questionnaire is freely available on the CHEXI website (www.chexi.se) in several different languages, including English, Swedish, French, Spanish, Chinese, and Farsi. It is meant to be used by parents or teachers, and it includes four a priori subscales measuring working memory (9 items), planning (4 items), inhibition (6 items), and regulation (5 items). Each item is rated on a scale from 1 (definitely not true) to 5 (definitely true), with higher scores indicating larger EF deficits. Please refer to Table 20.1 for the three statements loading highest on each of the four a priori subscales. The four a priori subscales were created based on Barkley’s (1997) hybrid model, in which inhibition, working memory, regulation, and planning are seen as constituting the major EF deficits in

| Table 20.1 Sample items from the four a priori subscales included in the CHEXI |
|---------------------------------|-------------------------------------------------------------|
| Working memory subscale         | Has difficulty understanding verbal instructions unless he/she is also shown how to do something |
|                                 | Has difficulty remembering what he/she is doing, in the middle of an activity |
|                                 | When asked to do several things, he/she only remembers the first or last |
| Planning subscale               | Has difficulty with task or activities that involve several steps |
|                                 | Has difficulty carrying out activities that require several steps (e.g., for younger children, getting completely dressed without reminders; for older children, doing all homework independently) |
|                                 | Has difficulty telling a story about something that has happened so that others may easily understand |
| Inhibition subscale             | Has difficulty holding back his/her activity despite being told to do so |
|                                 | Has difficulty stopping an activity immediately upon being told to do so. For example, he/she needs to jump a couple of extra times or play on the computer a little bit longer after being asked to stop |
|                                 | Gets overly excited when something special is going to happen (e.g., going on a field trip or to a party) |
| Regulation subscale             | Seldom seems to be able to motivate himself/herself to do something that he/she doesn’t want to do |
|                                 | When something needs to be done, he/she is often distracted by something more appealing |
|                                 | Has difficulty following through on less appealing tasks unless he/she is promised some type of reward for doing so |
children with ADHD. However, these four factors have not been identified through factor analyses of the CHEXI questionnaire in different age groups and in several different language versions, such as Swedish (Thorell & Nyberg, 2008), French (Catale, Lejeune, Merbah, & Meulemans, 2013; Catale, Meulemans, & Thorell, in press), Turkish (Kayhan, 2010), and Portuguese (Trevisan, Dias, Menezes, & Seabra, 2012). Instead, these studies have consistently identified two broad factors referred to as working memory (working memory and planning a priori subscales) and inhibition (inhibition and regulation a priori subscales). One exception is a Brazilian study (Trevisan et al., 2012), which found a two-factor solution for parent ratings but a one-factor solution for teacher ratings. The internal consistency of the two major factors of the CHEXI has proved to be satisfactory using both parent and teacher ratings. In addition, test-retest reliability for the CHEXI has been shown to be high (Catale et al., in press; Thorell & Nyberg, 2008) using parent ratings collected 3–10 weeks apart. Correlations between different raters have seldom been investigated using the CHEXI, although one study (Thorell, Veleiro, Siu, & Mohamadi, in press) reported modest, although significant, relations between parent and teacher ratings for both the inhibition ($r = .42$, $p < .001$) and working memory subscales ($r = .43$, $p < .001$).

Previous factor analyses have been conducted using children aged 5–7 years (Thorell & Nyberg, 2008), 5–6 years (Catale et al., 2013), 7–11 years (Catale et al., in press), 4–7 years (Trevisan et al., 2012), and 5–8 years (Kayhan, 2010). To our knowledge, no factor analysis of the CHEXI has been conducted using children below 4 or above 11 years. It should be noted that previous research using EF laboratory measures has demonstrated that it may not be fruitful to differentiate between inhibition and working memory in very young children (e.g., Wiebe, Espy, & Charak, 2008). On the other hand, a more differentiated EF profile has been found in older children, again using EF laboratory measures (e.g., Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; St. Clair-Thompson & Gathercole, 2006). In summary, the two-factor structure described above for children 4–11 years may not be valid for children outside this age range. However, it is also important to mention that the CHEXI was originally intended for use solely in the age range 4–12 years. If one is interested in EF deficits in children below 4 years, the preschool version of the BRIEF has been developed for children as young as 2 years, and more temperament-related measures of executive functioning (i.e., effortful control) can be made already in infancy using the Infant Behavior Checklist (Rothbart, 1981) or in preschoolers and early school-age children using the Child Behavior Checklist (Rothbart, Ahadi, Hershey, & Fisher, 2001). If one is interested in EF deficits in teenagers, we would recommend using rating scales with a greater emphasis on more complex executive skills such as organization, planning, and time management (e.g., the CEFI and the BDEFS-CA).

### Relations Between the CHEXI and EF Laboratory Tests

Several studies have examined the relation between ratings using the CHEXI and EF laboratory measures. The original study by Thorell and Nyberg (2008) found significant relations between the two CHEXI subscales and laboratory measures of inhibition (go/no-go task) and working memory (word span task). However, as can be seen in Table 20.1, the relations were modest, and both Catale et al. (in press) and Kayhan (2010) have failed to find any significant relations between the CHEXI and laboratory measures of executive functioning. Mixed findings with regard to this issue have also been presented for other EF rating instruments such as the BRIEF (Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Mahone et al., 2002; Vriezen & Pigott, 2002). It should be noted that despite the fact that laboratory measures are often regarded as the “golden standard” for measuring EF, the low correlations between ratings and laboratory measures should not necessarily be seen as a limitation. Instead, questionnaire ratings and laboratory measures most likely capture different
aspects of executive functioning. According to one hypothesis (Anderson et al., 2002), laboratory tests primarily capture the cognitive aspect of executive functioning, whereas rating instruments capture more emotional and social aspects. Another difference is that questionnaires provide reports on the child’s behavior in the “real world” and are based on observations made over an extended period of time, whereas laboratory tests are administered in a much more predictable and structured environment during a short period of time. In conclusion, EF tests and laboratory measures should be seen as complementary to one another, and a more complete picture of a child’s executive profile will be obtained by combining these two types of measures.

Clinical Utility of the CHEXI

Three previous studies have investigated the clinical utility of the CHEXI in ADHD populations. The first study (Thorell, Eninger, Brocki, & Bohlin, 2010) included 15 children with ADHD and 30 normally developing controls (age 7 years), and the results showed that the children in the ADHD group differed significantly from the comparison group on both the CHEXI inhibition and the working memory subscale, with large effect sizes (d = 1.79–2.95). In addition, a logistic regression analysis showed that, using either parent or teacher ratings, both the CHEXI inhibition and working memory scales contributed significantly to distinguishing between the ADHD group and the comparison group. The sensitivity (range .73–.93) and specificity (range .79–.93) were high for both parent and teacher ratings. The highest classification rate was obtained for parent ratings on the inhibition subscale, where 93.3% of the children were correctly classified.

The second study (Catale et al., in press), investigating the CHEXI in ADHD samples, included two subsamples of children aged 8–11 years from Belgium (25 ADHD children, 25 controls) and Sweden (62 ADHD children, 62 controls), and the results confirmed the findings of Thorell et al. (2010) by showing high sensitivity (range .84–.92) and high specificity (range .92–.96) for both the inhibition and working memory subscales in the two samples.

Third, another study compared three groups of children aged 5–7 years: (1) children diagnosed with ADHD, (2) children with executive and attention deficits (defined as having a Z-score below 1.5 on at least two executive and attention tasks assessing working memory, inhibition, cognitive flexibility or selective attention), and (3) normally developing children. The results showed that both clinical groups (ADHD group and EF deficits group) differed significantly from the comparison group on both the CHEXI inhibition and the working memory subscales (Catale, Lejeune, Merbah, & Meulemans, 2011). Furthermore, the CHEXI was shown to successfully distinguish between the respective clinical groups and the control group, with sensitivity and specificity ranging from .73 to .89.

Finally, in a non-clinical study from Belgium, the CHEXI was also examined in relation to other types of behavior problems. In this unpublished study on 80 normally developing children aged 8–11 years, associations were examined between the CHEXI and the Conners Parent Rating Scale (CPRS—48 items; Conners, 1970). Both rating scales were completed by the children’s parents. The CHEXI was found to be most strongly associated with learning disabilities (r = .73 for the working memory subscale and .75 for the inhibition subscale). Significant relations were also found to both the Hyperactivity/Impulsivity subscale (r = .27 for working memory and .48 for inhibition) and the Conduct Problem subscale (r = .27 for working memory and .39 for inhibition). Interestingly, no significant correlations were found between the CHEXI and the Psychosomatic and Anxiety subscales of the CPRS. Furthermore, the correlations between the CHEXI and the Conduct Problem subscale disappeared when controlling for hyperactivity/impulsivity. Thus, the significant relations between the CHEXI and conduct problems were caused by the overlap between conduct problems and ADHD symptoms. In conclusion, these findings suggest that the CHEXI subscales have good convergent and divergent validity.
The Relation Between the CHEXI and Functional Impairments

One of the criticisms of EF tests (e.g., Barkley & Fischer, 2011) has been that they have very low ecological validity (i.e., they are only weakly related to how EFs are used in daily life activities in a natural setting). An important question when using EF ratings is therefore how well a certain rating instrument correlates with functional impairments. As presented in Table 20.2, both the CHEXI inhibition and working memory subscales have been shown to be related to academic skills among children attending kindergarten (Thorell & Nyberg, 2008). Interestingly, the effects of the CHEXI working memory subscale were significantly related to early academic skills even when controlling for the effects of EF tests (see boldfaced figures in Table 20.2). Again, these results emphasize that ratings and tests capture at least partially different aspects of EF.

The relation between CHEXI and academic achievement has also been investigated in a cross-cultural study including children aged 5–12 years from four different countries: Sweden, Spain, Iran, and China (Thorell et al., in press). The results showed that both the inhibition and working memory subscales of the CHEXI were related to academic achievement (i.e., mathematics and language skills) in all four countries, with the exception of CHEXI parent ratings in China.

Finally, two recent longitudinal studies have examined the relation between CHEXI ratings and academic performance. The first is an unpublished Swedish study investigating CHEXI ratings in kindergarten (age 6) and school performance in grade 2 (age 8). This study found that teacher ratings on the CHEXI working memory subscale were significantly related to mathematics in grade 2, even when controlling for the effect of early mathematic abilities in kindergarten. Thus, the CHEXI working memory subscale was able to predict the change in academic achievement between the two time points. No significant relations were found for the CHEXI inhibition subscale or to other aspects of academic achievement (i.e., reading or writing), but it should be noted that this study had limited power to detect such relations given its small sample size (n=47). The second longitudinal study looked at a Spanish sample (Veleiro & Thorell, 2012) and found significant relations

<table>
<thead>
<tr>
<th>Laboratory measures</th>
<th>ADHD symptoms ( ^{ab} )</th>
<th>Early academic skills ( ^{b} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inhibition</td>
<td>Working memory</td>
</tr>
<tr>
<td>Parent ratings ((n=113))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working memory factor</td>
<td>.33***</td>
<td>.26**</td>
</tr>
<tr>
<td>Inhibition factor</td>
<td>.28**</td>
<td>.07</td>
</tr>
<tr>
<td>Teacher ratings ((n=89/105))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working memory factor</td>
<td>.29**</td>
<td>.39***</td>
</tr>
<tr>
<td>Inhibition factor</td>
<td>.35***</td>
<td>.19*</td>
</tr>
</tbody>
</table>

Reprinted by permission of the publisher (Taylor & Francis Ltd., http://www.tandf.co.uk/journals)

\(^{a}\) Relations to ADHD symptoms represent correlations across raters

\(^{b}\) Bold-faced entries denote relations that remained significant also when controlling for the effect of EF laboratory measures
between the CHEXI working memory subscale measured at 4 years and tests of basic mathematics abilities 12 months later.

In summary, there appears to be support for use of the CHEXI, especially the working memory subscale, as an early screening measure for early academic difficulties. To our knowledge, no other studies have presented data examining the relation between the CHEXI and other measures of functional impairments such as low prosocial skills, although there is at least one ongoing study investigating this issue.

Cross-Cultural Validation of the CHEXI in Different Countries

As stated above, the CHEXI is available in many different languages, and we are currently working on translating the questionnaire into even more languages (e.g., Danish, Dutch, and Japanese). These new versions will be available on the CHEXI website (www.chexi.se). As also stated above, the original factor structure of the instrument has been replicated in several different countries. A recent cross-cultural study, however, emphasized the need to take cultural biases into account when collecting ratings of problem behaviors in children using the CHEXI or any other rating instrument for that matter. As mentioned above, the cross-cultural study (Thorell et al., in press) included CHEXI ratings from Sweden, Spain, Iran, and China, and the results showed that there were significant effects of country with regard to both teacher and parent ratings on both CHEXI subscales. The main finding of the post hoc analyses was that the children in the Chinese sample received higher scores (i.e., showing higher EF deficits) compared to the children in the other countries. Interestingly, other studies using laboratory measures have shown the opposite pattern, with Chinese children being more skilled in executive functioning compared to children in the USA (e.g., Lan, Legare, Ponitz, Li, & Morrison, 2011; Sabbagh, Xu, Carlsson, Moses, & Lee, 2006). The authors of the cross-cultural study using the CHEXI therefore concluded that it is likely that the obtained cross-country differences do not reflect true differences in the children, but rather cultural biases. The Chinese culture has a strong emphasis on self-regulatory skills, and executive functioning deficits may therefore be exacerbated by strong cultural expectations. A similar conclusion was drawn by, for example, Hinshaw et al. (2011) with regard to ADHD symptoms.

Future Directions

To conclude, the CHEXI has good psychometric properties and can be considered a valuable screening instrument for identifying children at risk of developing ADHD, EF deficits, and early academic difficulties. However, as the instrument has not yet been nationally standardized in any country, it is at this time more valuable as a research tool than as a clinical instrument.

Finally, we also like to emphasize that there are many important avenues for future research on rating measures for executive functioning. Of most importance is perhaps to extend longitudinal investigations examining the development of executive functioning across early and middle childhood. Such an approach would allow us to gain further knowledge of the usefulness of EF rating instruments as an early screening measure for poor academic achievement and early behavior problems (primarily symptoms of ADHD). In addition, it would be valuable if future studies were to include several different EF rating instruments (e.g., CHEXI, BRIEF, and BDEFS-CA), allowing examination of both differences and commonalities. Finally, it should be mentioned that there is an adult version (i.e., self-rating instrument) of the CHEXI called the “Adult Executive Functioning Inventory” (ADEXI), which is still unpublished and thus far only available in English and Swedish. However, data from an ongoing study have shown that this instrument has good psychometric properties. It is for future studies to examine how well this adult version of the instrument can distinguish between adults with ADHD and controls and to examine its relation to EF laboratory tests.
References


