Developing and Validating a Short-Form Questionnaire for the Assessment of Seven Facets of Conscientiousness in Large-Scale Assessments

Patrick Franzen, A. Katrin Arens, Samuel Greiff, Lindie van der Westhuizen, Antoine Fischbach, Rachel Wollschläger & Christoph Niepel

View supplementary material

Published online: 17 Nov 2021.

Submit your article to this journal

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Developing and Validating a Short-Form Questionnaire for the Assessment of Seven Facets of Conscientiousness in Large-Scale Assessments

Patrick Franzen[1], A. Katrin Arens[2], Samuel Greiff[1], Lindie van der Westhuizen[1], Antoine Fischbach[3], Rachel Wollschläger[3], and Christoph Niepel[1]

[1]Department of Behavioral and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg; [2]DIPF | Leibniz Institute for Research and Information in Education, Frankfurt, Germany; [3]Luxembourg Centre for Educational Testing, University of Luxembourg, Esch-sur-Alzette, Luxembourg

ABSTRACT

Conscientiousness is the most important personality predictor of academic achievement. It consists of several lower order facets with differential relations to academic achievement. There is currently no short instrument assessing facets of conscientiousness in the educational context. Therefore, in the present multi-study report, we develop and validate a short-form questionnaire for the assessment of seven Conscientiousness facets, namely Industriousness, Perfectionism, Tidiness, Procrastination Refrainment, Control, Caution, and Task Planning. To this end, we examined multiple representative samples totaling $N = 14{,}604$ Grade 9 and 10 students from Luxembourg. The questionnaire was developed by adapting and shortening an existing scale using an exhaustive search algorithm. The algorithm was specified to select the best item combination based on model fit, reliability, and measurement invariance across the German and French language versions. The resulting instrument showed the expected factorial structure. The relations of the facets with personality constructs and academic achievement were in line with theoretical assumptions. Reliability was acceptable for all facets. Measurement invariance across language versions, gender, immigration status and cohort was established. We conclude that the presented questionnaire provides a short measurement of seven facets of Conscientiousness with valid and reliable scores.

Conscientiousness has been shown to be the strongest personality predictor of academic achievement, rivaling or even surpassing the predictive power of intelligence (Poropat, 2009). However, Conscientiousness has also been found to be a multidimensional construct with various lower-order facets being differentially related to indicators of academic achievement (de Vries et al., 2011; MacCann et al., 2009, 2015; Paunonen & Ashton, 2013; Rikoon et al., 2016). Hence, it is necessary to assess these lower-order facets when attempting to predict academic achievement.

In the past decades, the role of large-scale assessments in educational research has grown. One example of such a highly influential large-scale assessment is the well-known Programme for International Student Assessment (PISA; Organization for Economic Co-operation and Development, 2016). While large-scale assessments have certain benefits, they also pose additional challenges for researchers. One issue is the need to assess a multitude of variables while keeping the assessment time short. This problem is not only relevant for large-scale assessments, but also for many other assessment contexts. In particular, time limits can pose a

challenge when assessing personality constructs related to students' learning, as personality questionnaires usually include large numbers of items. To the best of our knowledge, there is no existing short-form questionnaire focusing on Conscientiousness as a multidimensional construct. The present paper aims to fill this gap by developing a short-form questionnaire to measure seven lower-order facets of Conscientiousness. We adapted and shortened an existing instrument by MacCann et al. (2009) using an exhaustive search algorithm. We relied on four different samples, including two fully representative cohorts of virtually all ninth-grade students in Luxembourg, which resulted in an overall sample of $N = 14{,}604$.

Conscientiousness and its multidimensional structure

Conscientiousness is a broad personality trait describing the tendency to be self-controlled, responsible, industrious, orderly, and rule-abiding (Roberts et al., 2009). It is part of the Big Five taxonomy of personality traits, which

encompasses Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness (see e.g., Digman, 1990 for an overview). Research interest in these personality traits as predictors of behavior is steadily increasing. Conscientiousness has been shown to predict behavioral outcomes in almost all aspects of life and throughout the lifespan (see e.g., Roberts et al., 2005), such as health and health behaviors (Bogg & Roberts, 2004; Hampson et al., 2007; Takahashi et al., 2013), life satisfaction (Hayes & Joseph, 2003; Smith et al., 2013), job performance (Dudley et al., 2006), and academic achievement (Kim et al., 2016; Poropat, 2009, 2014a; Richardson et al., 2012).

Conscientiousness consists of a number of different lower-order facets (DeYoung et al., 2007; Jackson et al., 2010; MacCann et al., 2009; Peabody & de Raad, 2002; Perugini & Gallucci, 1997; Roberts et al., 2004, 2005; Saucier & Ostendorf, 1999). The three most commonly identified facets are Industriousness, Orderliness and Self-Control. Industriousness is the tendency to work hard, work efficiently, strive for excellence, and exceed expectations. Orderliness describes the tendency to organize one's time and make plans as well as a tendency toward cleanliness and neatness. Self-Control comprises the ability to control one's impulses and reflect on one's behavior before acting. Other facets that have emerged in various empirical studies include Responsibility, Traditionality, Decisiveness, Formality, Punctuality, and Perseverance. Whereas there is no clear consensus on the validity of the latter facets, Industriousness, Orderliness and Self-Control are essential for the lower-order structure of Conscientiousness and have been consistently replicated (e.g., MacCann et al., 2009; Roberts et al., 2004, 2005).

Accounting for the lower-order facets is necessary when assessing Conscientiousness, as research indicates that they exhibit differential relations to various outcomes, including cognitive ability (Rikoon et al., 2016), and academic achievement (MacCann et al., 2009). For certain outcomes, such as grade-point-average (GPA), individual facets are better predictors than the broad Conscientiousness factor (MacCann et al., 2009; Paunonen & Ashton, 2001). Hence, assessment on the facet level offers incremental benefits over merely measuring the broad Conscientiousness factor.

Conscientiousness in educational research

Of all Big Five personality traits, Conscientiousness has been shown to be the strongest predictor of academic achievement as it demonstrates some of the highest associations with academic achievement ever reported (Poropat, 2009, 2014b; Richardson et al., 2012; see also Trautwein et al., 2015, Song et al., 2020). Moreover, Conscientiousness has been linked to several educationally relevant behavioral indicators of academic achievement, including lower class absenteeism (Chamorro-Premuzic & Furnham, 2003; Furnham et al., 2003; Lounsbury et al., 2004; MacCann et al., 2009), fewer rule violations (Ivcevic & Brackett, 2014), a lower school drop-out rate (Migali & Zucchelli, 2017), and better in-class behavior (Chamorro-Premuzic & Furnham, 2003; Furnham et al., 2003).

In light of the importance of Conscientiousness for academic achievement and motivation, it makes sense to take a closer look at its lower-order structure because lower-order facets of Conscientiousness might be differently related to outcomes of educational success. However, most attempts to uncover the different facets of Conscientiousness have been based on adult samples. This is problematic for educational research, as the lower-order facets found in adult samples might not generalize to adolescent samples. To the best of our knowledge, there is only one study investigating the lower-order structure of Conscientiousness in adolescents (MacCann et al., 2009), which was based on 117 Conscientiousness items from the International Personality Item Pool (IPIP; Goldberg et al., 2006). The items were rated by 291 high school students and then assessed using exploratory factor analysis. The study unveiled eight facets, namely *Industriousness*, *Perfectionism*, *Tidiness*, *Procrastination Refrainment*, *Control*, *Caution*, *Task Planning* and *Perseverance*. Perfectionism, Procrastination Refrainment and Industriousness reflect the content of the Industriousness facet reported in other studies. Perfectionism here describes the tendency to strive for perfection and outdo others. Procrastination Refrainment is the propensity to start tasks right away without wasting time or putting them off. Industriousness is defined as the tendency to work hard. MacCann et al. (2009) split the Orderliness facet described in previous research into Task Planning and Tidiness with Task Planning describing structuring one's time and making plans and Tidiness covering the neatness and cleanliness aspect. Self-control was separated into Control and Caution, with Control describing the tendency to control one's impulses and Caution describing the tendency to reflect before acting. The remaining facet—Perseverance—cannot be allocated to the three most commonly found facets (i.e., Industriousness, Orderliness and Self-control). The authors rather found that Perseverance is not located fully within the factor space of Conscientiousness, but instead overlaps with Neuroticism. Therefore, MacCann et al. (2009) concluded that Perseverance should be excluded when the goal is an assessment of pure Conscientiousness. This results in seven lower-order facets of pure Conscientiousness (i.e., Industriousness, Perfectionism, Tidiness, Procrastination Refrainment, Control, Caution, Task Planning), which are in line with the most commonly reported facets in previous research based on adult samples (Peabody & de Raad, 2002; Roberts et al., 2004, 2005).

Importantly, the facets described by MacCann et al. (2009) were also shown to have differential relations with academic outcomes. Industriousness appears to be the most important facet in this context. As such, Industriousness showed the highest correlation with GPA compared to the other facets and the broad Conscientiousness factor (MacCann et al., 2009, 2015; Rikoon et al., 2016). Industriousness describes the tendency to be hard working and ambitious—traits that will directly influence work ethic and thereby performance. Hence, it is only logical that previous research found it to be the best predictor of GPA,

driving a large portion of the predictive power of general Conscientiousness (see also MacCann et al., 2015; Roberts et al., 2005). Industriousness has further been shown to predict several behavioral indicators of academic achievement, such as class absence (MacCann et al., 2009). In contrast, Tidiness has so far been shown to have almost no predictive utility for virtually any academic outcomes, including GPA, teacher ratings, and behavioral indicators (MacCann et al., 2009; Rikoon et al., 2016). All other Conscientiousness facets have shown at least some predictive utility for GPA (MacCann et al., 2009, 2015; Rikoon et al., 2016).

In addition, no empirical study so far has examined the role of the different facets of Conscientiousness as correlates for standardized achievement test scores (SATS). As SATS are one of the most widely used indicators of academic achievement, studying the relations between SATS and the different facets of Conscientiousness is crucial for evaluating the utility of the Conscientiousness facets in education. Previous evidence has shown a smaller association between the broad Conscientiousness construct and SATS compared to GPA (Noftle & Robins, 2007). This makes sense considering that GPA includes teacher ratings of students' behavior, which have been shown to be linked to Conscientiousness (MacCann et al., 2009). Hence, one might assume that the separate facets of Conscientiousness also exhibit lower associations with SATS compared to GPA. However, Caution— as a lower-order facet of Conscientiousness—might show a divergent pattern showing higher relations with SATS compared to GPA. Caution has been shown to be positively related to intelligence (Rikoon et al., 2016)—probably due to the linkages to inhibitory processes (e.g., Dempster, 1991); and intelligence has also been shown to be positively associated with higher SATS (see, e.g., Frey & Detterman, 2004; Koenig et al., 2008). In addition, constructs closely related to Caution—for example self-control—have been shown to also be related to SATS (Duckworth et al., 2012; Meldrum et al., 2017). Students displaying higher self-control might be able to focus on homework and paying attention in class more easily, thereby gaining an academic advantage and ultimately higher SATS (see e.g., Duckworth et al., 2012). Therefore, it can be assumed that Caution might be a useful predictor of SATS, likely exhibiting a higher relation to SATS compared to the other facets of Conscientiousness and the broad Conscientiousness factor.

The need for and development of short instruments

In light of the importance of the facets of Conscientiousness for educational outcomes, there is a need for psychometrically sound instruments for their assessment in the educational context. Along with their framework, MacCann et al. (2009) presented a Concise Conscientiousness Measure (CCM) for the assessment of eight facets of Conscientiousness (or seven pure Conscientiousness facets, excluding Perseverance). The questionnaire is tailored toward use with adolescent students, as it was developed using a high school student sample. However, the seven factor-pure facets of Conscientiousness in the questionnaire

contain 59 items. This large number of items limits the applicability of the questionnaire to contexts where researchers have sufficient time to fit a large number of items into their assessment battery. The use of long questionnaires is often uneconomic, as it involves a large investment of both time and money. Hence, there is a need for a short-form questionnaire assessing Conscientiousness facets.

Most short scales are developed from existing longer questionnaires by successively deleting items, based on indicators such as Cronbach's alpha after deleting the item or the item's correlation with the overall scale. While comparably simple, these traditional procedures often fail to account for multiple psychometric criteria, resulting in an unsatisfactory factor structure, insufficient reliability, or lack of invariance, amongst others (Olaru et al., 2015). Unsurprisingly, many short scales have received criticism for their overall psychometric properties (Emons et al., 2007; Kruyen et al., 2013). Thus, a different approach is required to develop psychometrically sound short scales. Recently, algorithmic approaches to shortening questionnaires have demonstrated promising results (Olaru et al., 2015; Schroeders et al., 2016). These algorithmic approaches offer the possibility to optimize short scales with respect to any combination of (psychometric) criteria. Algorithms that have been applied to short scale development thus far, such as ant-colony optimization (e.g., Leite et al., 2008) and the genetic algorithm (e.g., Yarkoni, 2010), can be classified as heuristic algorithms. Heuristic algorithms aim to find a near-optimal solution using specific search patterns, when testing all possible solutions is not possible due to computation time. As a tradeoff, heuristic algorithms are not guaranteed to find the best solution, and may return different solutions when run multiple times on the same data. Nonetheless, they have shown to be extremely useful for the development of short scales, easily outperforming traditional psychometric approaches (Olaru et al., 2015; Schroeders et al., 2016). If testing all possible combinations of items is possible, one may instead rely on an exhaustive search algorithm (ESA). The ESA systematically enumerates all possible solutions to a given problem (i.e., all possible item combinations from a larger item pool), then tests each of these solutions against certain criteria defined a priori, ultimately returning the single best solution. Thereby, the ESA is guaranteed to find the best item combination based on the criteria specified a priori for the underlying data. As a tradeoff, the ESA may require long computation times. With larger amounts of input (e.g., more items in a questionnaire), computation times can become unfeasible, easily surpassing the researchers' lifetime. However, when computation time is not an issue, an ESA should be preferred over a heuristic algorithm.

The present article

The importance of Conscientiousness and its lower-order facets creates a need for high-quality assessment instruments. In light of the growing need for short scales, such an instrument should be as parsimonious as possible, while still

being psychometrically sound. Currently, there is no short questionnaire assessing the lower-order facets of Conscientiousness. Hence, in the present research, we develop and validate a short-form instrument for the assessment of seven pure Conscientiousness facets using an ESA approach. We specifically tailored the questionnaire to assessment with adolescents in (large-scale) educational studies. Specifically, we conducted three separate studies to develop and validate a short form of the CCM (MacCann et al., 2009), which we term the CCM-S. We made use of two entire cohorts of ninth-grade students in Luxembourg in 2017 and 2018, as well as additional smaller samples from 2019. It should be noted that Luxembourg has an official multilingual educational policy with two languages of instruction in secondary school: French and German. Therefore, Study 1 describes the development of the German and French language versions of the CCM-S using an ESA algorithm. Study 2 assesses psychometric properties of the CCM-S, such as factorial and criterion validity (with SATS as criterion), reliability as well as measurement invariance across the French- and German-language versions, gender, immigration status, and different cohorts. In Study 3, we assess the convergent, discriminant, and criterion validity of the CCM-S by relating its facets to theoretically overlapping constructs, the Big Five and GPA. The studies were not preregistered.

Study 1

Study 1 aimed to shorten the CCM presented by MacCann et al. (2009). As our goal was an assessment of pure Conscientiousness, Perseverance was excluded. We decided to select four items for each of the remaining seven facets by means of an ESA, resulting in a seven-facet, 28-item questionnaire. Four items ensure overidentification of latent models when modeling individual facets. Latent modeling of individual facets was considered important to make sure that each facet could be assessed and used on its own, as is often done when developing short scales (see e.g., Niepel et al., 2019, Soto & John, 2017b). The main goal during development was to optimize factorial validity, reliability as well as measurement invariance across the German and French language versions of all individual facet scales.

Methods

Sample and procedure
The sample for Study 1 was the entire population of ninth grade students in Luxembourg in November and December 2017, which came to a total of $N = 6{,}235$ students (47.9% female), with a mean self-reported age of 15.01 years ($SD = 1.03$, range 12–25). Students were clustered into 351 classes in 34 schools. The number of missing values ranged between 9% and 25.1% per item.

The data used in the present study derive from the Luxembourgish school monitoring programme, Épreuves Standardisées (ÉpStan; Martin et al., 2014), which assess full cohorts of students in Luxembourgish public schools in Grades 1, 3, 5, 7, and 9 at the beginning of each school year. The ÉpStan are prepared and organized by Luxembourg Center for Educational Testing and conducted during regular school hours by class teachers following a standardized procedure with a fixed assessment time. The ÉpStan assess students' academic achievement in different domains depending on grade level, learning motivation including academic interest, academic self-concept (i.e., students' mental representation of their ability), and attitudes toward school. In secondary school, that is, from Grades 7 and 9, the ÉpStan are web-based, with Grade 7 being assessed on tablets and Grade 9 on computers. For the cohort of ninth grade students in 2017, we included the 59-item version of the CCM (as described in the next paragraph) into the ÉpStan assessment battery. As ninth grade students have a mixture of German and French as the language of instruction, depending on the subject and their academic track, German- and French-language versions of all questionnaires were implemented, and students could freely switch between the two languages at all times.

Concise conscientiousness measure (CCM)
The CCM presented by MacCann et al. (2009) comprises eight facets of Conscientiousness, namely *Industriousness, Perfectionism, Tidiness, Procrastination Refrainment, Control, Caution, Task Planning, and Perseverance* (see introduction section for full construct definitions). It was developed using a U.S.-based adolescent sample consisting of 13–19 year olds ($N = 291$), and constructed using a pool of 117 English-language items from the IPIP (Goldberg et al., 2006). The IPIP is a free and open source repository for personality items in over 25 languages. In the original investigation by MacCann et al. (2009), these items were first investigated using parallel analysis, followed by an exploratory factor analysis (EFA), and finally a series of confirmatory factor analyses (CFA). For the latter, a one-factor solution and an eight-factor solution were tested. The eight-factor solution demonstrated better fit and salient loadings for all items but one. Cronbach's α for all facets ranged from .80 (Control and Caution) to .91 (Industriousness). Correlations of the facets with other Conscientiousness measures ranged from .41 (Perfectionism) to .72 (Procrastination Refrainment). Including all eight facets, the CCM contains a total of 68 items. As described earlier, we excluded Perseverance to retain only the seven pure Conscientiousness facets, resulting in a total of 59 items.

For the present investigation, the English items were first translated into German and French by bilingual experts from the Luxembourg Center for Educational Testing following the team approach described by Behr et al. (2016), which involves multiple experts and a multi-stage process. The translated items were not pretested. The German and French wording of the final CCM-S (as well as the corresponding English wording) can be found in the Appendix. Students were able to switch between language versions for each item. Negatively worded items were reverse-coded so that higher values represented stronger expression of the criterion. All items were answered on a five-point Likert scale. Reliabilities for the seven subscales of the CCM were

acceptable to good in the present study, ranging from $\omega =$ .76 (Industriousness) to $\omega = .87$ (Task Planning). The only exceptions were Procrastination Refrainment and Tidiness, at $\omega = .52$ and $\omega = .67$ respectively.

Item selection procedure and data analysis

To get a better understanding of the full CCM as used in our sample, we estimated a CFA model of the 59-item instrument. A seven-factor model in which the items of each facet loaded only onto their respective factor was specified. All correlations between factors were freely estimated. The nested data structure was accounted for using the TYPE = COMPLEX command in Mplus 8 (Muthén & Muthén, 1998/2007), and missing data were handled by using full information maximum likelihood (FIML) with robust maximum-likelihood (MLR) estimation. Estimates from FIML have been shown to be valid and reliable for data missing (completely) at random (Enders, 2010; Graham, 2009).

Regarding our algorithmic procedure, we specified the ESA to evaluate all possible combinations of four items for each facet (see online supplemental material for the full algorithm syntax). The algorithm was designed to first gather all possible combinations of four items out of the original item pool for each facet. As the original CCM facets consisted of seven to ten items, this resulted in 35–210 possible four-item combinations for each facet. Then, the algorithm assessed all item combinations with respect to multiple criteria simultaneously, and finally returned the item combination that best met all criteria combined. We chose three relevant criteria. The first criterion was model fit as measured by CFI and RMSEA, with cutoffs of .95 and .05, respectively, the same cutoffs applied in previous algorithmic approaches to item selection (e.g., Schroeders et al., 2016). The second criterion was reliability as measured by McDonald's ω (McDonald, 1999). While there is no gold-standard cutoff value for McDonald's ω, we applied the commonly used cutoff of .70 indicating acceptable reliability (Dunn et al., 2014), which was also used in previous studies applying algorithmic item selection (Schroeders et al., 2016). The third criterion was measurement invariance between the German- and French-language versions. Measurement invariance is necessary for a fair scale, as it ensures that differences in means and variances between groups are not due to the measurement instrument itself, thereby enabling meaningful comparisons across groups. The most common approach is to assess different levels of invariance by comparing their respective model fits. Configural invariance ensures equal factor structure, metric invariance ensures equal factor loadings, and scalar invariance further ensures equal item intercepts across groups (Greiff & Scherer, 2018). According to recommendations by Cheung and Rensvold (2002) and Chen (2007), values smaller than or equal to ΔCFI $= -.01$ indicate invariance. To keep the algorithm parsimonious, we directly assessed the ΔCFI between the configural (equal factor structure) and scalar (equal loadings and equal intercepts) model (Putnick & Bornstein, 2016). For each of the criteria specified a priori, the algorithm

assigned an individual search weight to each item combination. The search weights were calculated using the logit transformation approach described by Schroeders et al. (2016) and Janssen et al. (2017). The benefit of this approach is that it increases the algorithm's differentiation around the relevant criterion's cutoff value. Furthermore, this approach standardizes search weights to range from zero to one, making the different criteria more easily comparable. For example, a CFI of .95 was transformed to an individual search weight of .50.

$$\varphi\text{CFI} = \frac{1}{1 + e^{95-100\text{CFI}}}.$$

Thus, a difference in CFI of .01 around the .95 cutoff was more impactful than a difference of .01 around a CFI value of .80, for example, thus making meaningful differences more impactful for item selection. This approach was applied to all selection criteria. For RMSEA, .05 was transformed into a search weight of .50. As smaller numbers indicate better values for RMSEA, the result was then subtracted from one.

$$\varphi\text{RMSEA} = 1 - \frac{1}{1 + e^{5-100\text{RMSEA}}}.$$

CFI and RMSEA were averaged into one individual search weight for model fit, following the approach by Schroeders et al. (2016).

$$\varphi\text{Fit} = \frac{\varphi\text{CFI} + \varphi\text{RMSEA}}{2}.$$

For reliability, values around the cutoff value of $\omega = .70$ were transformed into a search weight of .50.

$$\varphi\text{Rel} = \frac{1}{1 + e^{7-10\text{Rel}}}.$$

Finally, we used the same approach for measurement invariance. Hence, values of ΔCFI $= -.01$ were transformed into a search weight of .50. Here again, as smaller values indicate stronger invariance, the result was subtracted from one.

$$\varphi\text{MI} = 1 - \frac{1}{1 + e^{5-500\Delta\text{CFI}}},$$

with

$$\Delta\text{CFI} = \mid \text{CFI}_{\text{scalar}} - \text{CFI}_{\text{configural}} \mid.$$

The three individual search weights, namely model fit, reliability, and measurement invariance were then summed up with equal weighting to form a final search weight, so that each item combination was associated with one final search weight.

$$\max f(x) = \varphi\text{Fit} + \varphi\text{MI} + \varphi\text{Rel}.$$

The best possible item combination was determined by comparing these final search weights. The algorithm was written in R version 3.4.2 (R Core Team, 2007). To calculate the different criteria, the algorithm was configured to specify and run CFAs in Mplus 8 (Muthén & Muthén, 1998/2007). The resulting best item combination for each facet was checked by the authors to ensure that the content coverage

of the algorithm-selected items for each facet was sufficient. As all selected item combinations exhibited satisfactory content coverage, they were taken as the final CCM-S facet scale without any further manipulation. After item selection, the correlations between all individual CCM-S facets and the corresponding CCM facets were assessed using Pearson correlations. Higher correlations were regarded as more desirable, as the goal was to stay close to the content structure of the original scale despite using fewer items.

After this item selection process, we randomly split our sample in two halves, and ran the ESA in both of these subsamples.[1] This was done to ensure that the final item selection could be replicated, and was not solely based on over-optimization toward our specific sample. We then compared the results of the full sample (hereafter referred to as "final selection") to the results found in these subsamples. One subsample showed an almost identical item selection, only one facet had one different item selected. In the second subsample, four facets had one different item selected, while for the other three facets the item selection was identical to our final selection. As the results of these subsamples and our full sample were highly similar, we conclude that our final selection was not based on over-optimization toward the overall sample.

Measurement invariance tests after the item selection followed the common approach of testing each individual level of invariance separately. Whenever full metric or scalar invariance was not given, partial invariance was tested by freeing one parameter based on modification indices (Greiff & Scherer, 2018).

Results

Regarding the full CCM, the overall CFA model did not show an acceptable fit: χ^2 (1631) = 34095.77 (p < .001), CFI = .65, RMSEA = .06, SRMR = .013. All factor loadings are reported in the online supplemental material.[2]

The ESA yielded a questionnaire consisting of seven individual four-item scales. The German and French wording of the final CCM-S items (as well as the corresponding English wording) can be found in the Appendix. The final four-item scales exhibited good model fit, reliability, and scalar measurement invariance across the German and French language versions, except for Industriousness and Procrastination refrainment, which exhibited only partial scalar invariance with one intercept freed, and Tidiness, which exhibited partial metric and scalar invariance with one loading freed (Table 1). All CCM-S facets showed high correlations with their counterpart in the full CCM, ranging between r = .83 and r = .93, ps < .001 (Table 1).

[1]This process was suggested by an anonymous reviewer and was therefore done after the final item selection. For future investigations, we highly suggest using a similar approach, as it offers a quick and effective way to avoid over-optimization, and adds to the validity of the final item selection.
[2]This model did not converge properly without post-hoc model adjustments. As this model was only estimated for comparison purposes, and was not used for any further analyses, we did not integrate any post-hoc model modifications but report this initial model.

Table 1. Final values of the CCM-S scales on all selection criteria and correlations with the respective original facet scale in the development sample.

Facet	CFI	RMSEA	ω	r	Invariance
Industriousness	1.00	.01	.80	.87***	Partial Scalar
Perfectionism	1.00	.01	.74	.92***	Scalar
Procrastination Refrainment	0.99	.04	.62	.90***	Partial Scalar
Control	1.00	.03	.64	.88***	Scalar
Caution	1.00	.00	.73	.86***	Scalar
Task Planning	0.99	.04	.79	.93***	Scalar
Tidiness	1.00	.02	.69	.83***	Partial Metric and Scalar

Note. ΔCFI = difference between the configural and scalar model when assessing measurement invariance of the German- and French-language versions, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation, ω = McDonald's ω, r = correlation between the CCM-S scale selection and the corresponding facet of the original CCM. *** = p < .001.

Discussion

In Study 1, we developed a short form of the CCM (MacCann et al., 2009), the CCM-S. The questionnaire was reduced from 59 to 28 items. The CCM-S covers seven facets of Conscientiousness and is available in German and French. We aimed to provide evidence for the factorial validity, reliability, and measurement invariance of the CCM-S across language versions. The initial assessment of the facets within the development sample was promising, as all facets returned by the algorithm exhibited good values on all selection criteria, with only minor exceptions. Furthermore, all facets of the CCM-S were very strongly correlated with the corresponding facets of the original CCM, indicating content validity. Study 2 was conducted with the aim of providing evidence for the validity of the CCM-S using a second independent sample.

Study 2

The goal of Study 2 was to provide further evidence for the validity of the CCM-S developed in Study 1. Our goal was to examine factorial validity, reliability, and measurement invariance for the overall seven-factor model and each individual facet. In addition, we aimed to assess criterion validity via the correlation between each facet and SATS in German, French, and math. In line with previous research demonstrating associations between Caution and Intelligence, and in turn Intelligence and SATS, we expected Caution to show the strongest association with SATS across all subjects (Frey & Detterman, 2004; Koenig et al., 2008; Rikoon et al., 2016).

Methods

Sample and procedure

Two samples from the ÉpStan were used for this study. The first sample (hereafter referred to as the 2018 sample) consisted of the whole population of ninth grade students in Luxembourg in November and December 2018, N = 6,279 students (47.8% female), clustered into 342 classes in 34 schools. Students' self-reported mean age was 15.03 years (SD = 1.06; range: 13–25 years). Missing values ranged from 11.3% to 25.6% on the item level. The second sample

(hereafter referred to as the 2019 sample) consisted of $N = 1,670$ ninth grade students in Luxembourg in November and December 2019 (47.8% female; age), corresponding to about a quarter of all ninth-grade students. Self-reported mean age was 14.9 years ($SD = 1.1$, range: 13–24 years). Students were clustered into 175 classes in 34 schools. Missing values ranged from 14.37% to 35.93% on the item level. The 2019 sample was only used to test measurement invariance across the 2018 and 2019 student cohorts. This was done to provide incremental evidence for the generalizability and applicability of the CCM-S across different samples by demonstrating that its factor structure replicates across different cohorts. For all other analyses, the 2018 sample was used.

Measures

Standardized achievement test scores (SATS).
Our measurement of SATS comprised German reading comprehension, French reading comprehension as well as mathematics and was assessed within the context of the ÉpStan (Martin et al., 2014). The tests were developed in accordance with the Luxembourgish national curriculum to provide feedback on educational outcomes to students, teachers, and the Luxembourg Ministry of Education. The tests were validated and pretested before administration to ensure that they accurately measure the competency standards defined by the Luxembourg Ministry of Education. In Grade 9 (the grade level analyzed in this study), three different test versions with varying difficulty levels are available depending on the academic track students are enrolled in. Nevertheless, as test scores are scaled by means of a unidimensional Rasch model, academic achievement can be compared across the different test versions.

Immigration status.
Immigration status is operationalized through students' and their parents' country of birth. Students are classified as native, if they and at least one of their parents were born in Luxembourg. First-generation immigration status is characterized by students being born outside of Luxembourg regardless of their parents' country of birth. Second-generation immigration status is defined by students being born inside and both parents being born outside of Luxembourg.

Data analysis

To investigate the factorial validity of the CCM-S, we calculated CFAs using Mplus 8. We first calculated a seven-factor model in which all four items of each facet loaded only onto their respective factor. All correlations between factors were estimated. Afterwards, we calculated separate CFAs for each facet. As in Study 1, the TYPE = COMPLEX command in Mplus was used to account for the nested data structure, nonnormality was accounted for by using the MLR estimator with FIML to account for missing values, and model fit was assessed by evaluating CFI, RMSEA, and SRMR based upon Hu and Bentler (1999) recommendations. Accordingly, CFI >.95, RMSEA <.05 and SRMR <.08 were considered

Table 2. Model fit and reliabilities for the individual facet models in the validation sample.

Facet model	df	χ^2	χ^2 p value	CFI	RMSEA	SRMR	α	ω
Industriousness	2	124.63	<.001	0.98	.11	.02	.86	.86
Perfectionism	2	17.20	<.001	1.00	.03	.01	.79	.79
Procrastination Refrainment	2	13.77	<.010	1.00	.03	.01	.80	.80
Control	2	336.01	<.001	0.88	.18	.05	.69	.71
Caution	2	275.96	<.001	0.95	.16	.03	.85	.85
Task Planning	2	70.62	<.001	0.99	.08	.02	.85	.86
Tidiness	2	38.91	<.001	0.99	.06	.02	.78	.80

Note. Df = degrees of freedom, χ^2 = value for chi-square test of model fit, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Square Residual, α = Cronbach's α, ω = McDonald's ω.

good fit, while CFI >.90, RMSEA <.08 were considered indicative of acceptable fit, in line with common practices.

Reliability was assessed with McDonald's ω (McDonald, 1999) and the commonly used Cronbach's α. For both reliability estimates, values over .70 were considered acceptable.

Measurement invariance was assessed by comparing nested models in terms of ΔCFI and ΔRMSEA. For each facet, we first tested configural invariance followed by metric invariance. If the latter was given, we further assessed scalar invariance. For the seven-factor model, factor correlation invariance was then tested if the previous levels could be established. Following Chen's (2007) recommendations, ΔCFI < −.01 and ΔRMSEA < .015 were considered indicative of invariance at each level. We tested invariance across gender (young men: $n = 3,251$; young women: $n = 3,000$), language version (French $n = 1,261$; German, $n = 4,282$), immigration status (native, $n = 2,692$; first generation, $n = 1,559$; second generation, $n = 1992$), and cohort (2018, $n = 6,279$; 2019, $n = 1,670$). Whenever full invariance was not given, partial invariance was tested by freeing one parameter based on modification indices.

Criterion validity was assessed by calculating the correlations between the CCM-S facet mean scores and SATS in math, German reading comprehension, and French reading comprehension. As we expected Caution to show a higher correlation with SATS than all other facets, we used the Fisher r-to-z transformation to assess significance of differences (Eid et al., 2015).

Results

The overall seven-factor model including all individual facets showed acceptable fit to the data, χ^2 (329) = 3965.09 ($p <$ .001), CFI = .93, RMSEA = .04, SRMR = .05. In terms of individual scales, Perfectionism, Procrastination Refrainment, Task Planning, and Tidiness showed acceptable to good fit. Industriousness and Caution showed acceptable fit for all values except RMSEA. Only the Control scale exhibited less than acceptable fit in general (Table 2). All facets provided reliable scores according to McDonald's ω (Table 2). In terms of Cronbach's alpha, scores for all facets except Control indicated acceptable reliability (see Table 2).

Regarding gender invariance, the overall seven-factor model exhibited scalar measurement invariance and was invariant on the factor correlation level. Of the individual

facet models, Procrastination Refrainment, Task Planning, and Tidiness exhibited scalar invariance. Industriousness, Perfectionism, Control and Caution exhibited partial scalar invariance across gender, with one intercept freed (Items 3, 16, 9, and 5, respectively, see Appendix). Concerning invariance across the German- and French-language versions, the overall model exhibited scalar invariance and invariance on the factor correlation level. All of the individual facet scales were invariant on the scalar level. Regarding immigration status, the overall model exhibited scalar invariance as well as invariance on the factor correlation level across all three groups: natives, first-generation immigrants, and second-generation immigrants. Of the individual facet scales, Perfectionism, Procrastination Refrainment, Caution, Task Planning, and Tidiness were invariant on the scalar level across all three groups. Industriousness exhibited scalar invariance between first- and second-generation immigrants, and partial scalar invariance between natives and first-generation immigrants as well as between natives and second-generation immigrants, with one intercept freely estimated (Item 2, see Appendix). Control exhibited scalar invariance between second- and first-generation immigrants and between second-generation immigrants and natives, and partial scalar invariance between natives and first-generation immigrants, with one intercept freely estimated (Item 10, see Appendix). Finally, regarding invariance across the 2018 and 2019 student cohorts, the overall model exhibited scalar invariance as well as invariance on the factor correlation level. All of the individual facet scales were invariant on the scalar level across student cohorts.

When examining criterion validity, Industriousness, Perfectionism, Task Planning, and Caution were associated with higher SATS in math, German reading and French reading. As expected, Caution correlated significantly with SATS in Math ($r = .13$, $p < .001$), German reading comprehension ($r = .21$, $p < .001$) and French reading comprehension ($r = .17$, $p < .001$). Tidiness, Procrastination refrainment and control showed little to no association with SATS across all subjects (Table 3). The correlation of Caution with SATS in math, French reading comprehension and German reading comprehension was significantly higher than all other facets at the $p < .001$ level for all comparisons. The same was true for the comparison of Caution with overall Conscientiousness regarding SATS in math ($p < .001$), German reading ($p < .001$), and French reading ($p = .002$)

Discussion

The goal of this study was to provide evidence of the validity of the CCM-S by examining its factorial validity, reliability, criterion validity, and invariance across different student subgroups (defined by gender, immigration status, cohort, and language version). This was achieved for both the overall questionnaire as well as all separate facet scales. The sole exception was the Control scale, as its model fit was below acceptable.

These findings have important implications for the future use of the CCM-S. Researchers wanting to implement only

Table 3. Correlation of individual facets and overall Conscientiousness with indicators of academic achievement.

Facet	SATS Math	SATS German	SATS French	GPA
Industriousness	-.03*	-.07*	-.10*	.34***
Caution	-.13***	-.21***	-.17***	.28***
Control	-.02	-.01	-.03*	.19***
Perfectionism	-.07***	-.08***	-.09***	.36***
Procrastination Refrainment	−.02	−.01	−.01	.19***
Task Planning	-.04**	-.05***	-.11***	.20***
Tidiness	−.01	−.02	-.04*	.05
Mean Conscientiousness	-.06**	.08**	.12**	.34***

Note. SATS = Standardized achievement test scores, GPA = Grade point average. SATS were assessed in Study 2; GPA was assessed in Study 3 using a different sample. *$p < .05$, **$p < .01$, ***$p < 001$, no indication $p > .05$.

certain Conscientiousness facets in their research may do so without fear of sacrificing psychometric quality for a shorter assessment time (with the exception of the Control facet), as the psychometric assessment was performed for each individual facet. As our results confirm measurement invariance, results from subgroups defined by gender, language, and immigration status can be compared. Furthermore, by demonstrating invariance across student cohorts, we provide additional evidence for the generalizability of our results across student samples.

The facet scales showed differential relationships to academic achievement. In line with our assumptions, Caution had a moderate association with SATS, which was significantly higher than the association of all other facets with SATS. In addition, Industriousness, Perfectionism, and Task Planning exhibited significant yet small associations with SATS in German, French and math, which is in line with research on the general association between Conscientiousness and SATS (Noftle & Robins, 2007). It has to be noted that due to our large sample size, even very small correlations were statistically significant, which is not necessarily indicative of practical relevance. The fact that most facets showed either no or only small correlations with SATS is in line with our expectations as well as previous results on the relation between Conscientiousness and SATS (Noftle & Robins, 2007).

While the results of this study provided evidence for the psychometric soundness of the CCM-S with respect to multiple criteria, we did not assess the relations between the individual facets with GPA. As the relation between facets of Conscientiousness and GPA has been assessed in previous studies (MacCann et al., 2009; Rikoon et al., 2016), demonstrating similar relations between the CCM-S facets and GPA would provide further evidence of the criterion validity of the scales. Furthermore, we did not yet assess convergent and discriminant validity of the CCM-S. These latter criteria are of major importance, as item selection for the facet scales of the CCM-S was based purely on statistical criteria (see Study 1). Thus, a full investigation of convergent and discriminant validity is required.

Study 3

After providing evidence for the validity of the CCM-S in Study 2, we aimed to extend this validation by covering aspects that remained unexamined. Therefore, in Study 3,

we assessed the convergent and discriminant validity of the CCM-S. We further investigated the criterion validity of the CCM-S using GPA as an outcome measure. We expected all facets except Tidiness to show significant positive associations with GPA based on previous evidence (MacCann et al., 2009). Industriousness was expected to show the strongest association with GPA (MacCann et al., 2009; Rikoon et al., 2016). As previous research on the relations between facets of conscientiousness and GPA has found correlations of around $r = .20$ (MacCann et al., 2009; Rikoon et al., 2016), we conducted a power analysis using $r = .20$ as the estimated effect size, a power of .80, and an alpha level of .05. The power analysis revealed a minimum sample size of $N = 153$.

Methods

Sample and procedure

Study 3 relied on an overall sample of $N = 330$ 10th grade students in Luxembourg in February 2019 from a total of 17 classes and 9 different schools in the academic track.[3] Age, gender, immigration background and other personal information was not collected. Missing values ranged from 0% to 11.82% on the item level. In contrast to the samples used in the previously described studies, this sample was not assessed as part of the ÉpStan. However, the data collection was organized and conducted in collaboration with the ÉpStan team at the Luxembourg Center for Educational Testing. The assessment took place in classrooms during regular school hours using laptops or tablets. One researcher was present throughout the entire assessment. Students received no remuneration for participation. All students and their legal guardians provided informed consent. The students were informed that they were free to stop their participation at any point without consequences.

Measures

Big Five inventory 2 short version (BFI-2-S).
The German version (Danner et al., 2016) of the BFI-2-S (Soto & John, 2017b), the short form of the BFI-2 (Soto & John, 2017a), was used to measure the Big Five personality constructs. Each construct consists of three facets, and each facet contains two items. However, the facet structure was disregarded in the present study, as the facets should only be used with large numbers of participants (Soto & John, 2017b). Instead, only the overarching Big Five constructs were considered. Reliabilities reported by the authors range from $\alpha = .73$ to $\alpha = .84$. For the present study, the German items were further translated to French by bilingual experts from the Luxembourg Center for Educational Testing. All items were answered on a five-point Likert scale. It should be noted that the full-length BFI-2 and its short form, the BFI-2-S, which was used here, were developed with an adult sample and have not yet been validated in an adolescent

sample (Soto & John, 2017a), which may decrease reliabilities (Rammstedt & Farmer, 2013; Soto et al., 2008). In line with this, reliabilities for Openness to Experience, Agreeableness and Neuroticism were below $\omega = .60$ in the present study. Thus, the results for these scales need to be interpreted with caution, but are nevertheless reported for the sake of completeness (see Table 4). English example items include "I am someone who …": "Tends to be quiet" (Extraversion), "Is compassionate, has a soft heart" (Agreeableness), "Is reliable, can always be counted on." (Conscientiousness), "Worries a lot." (Negative Emotionality), and "Is fascinated by art, music, or literature." (Open Mindedness).

NEO-personality inventory-revised conscientiousness scale.
The Neo-PI-R (Costa & McCrae, 1992) consists of eight scales measuring different facets of Conscientiousness, defined as Competence, Order, Dutifulness, Achievement Striving, Self-Discipline, and Deliberation, which are each measured with eight items. All 48 items were answered on a five-point Likert scale. Published and commercially available German and French versions of the NEO-PI-R Conscientiousness scale were used. The reported reliabilities for each facet range from $\alpha = .44$ to $\alpha = .84$ in an adolescent sample (Costa & McCrae, 1992). In our sample, reliability estimates ranged from $\omega = .61$ (Dutifulness) to $\omega = .81$ (Self-Discipline; see Table 4). Example items include (translated from the German version): "I think twice before answering a question" (Deliberation) and "I work hard to achieve my goals" (Achievement Striving).

The instrument was used to assess convergent and discriminant validity for the different CCM-S facets, following the approach by Rikoon et al. (2016) to assess the convergent and discriminant validity of the CCM. More specifically, the NEO-PI-R Order facet was assumed to be analogous to the Tidiness and Task Planning facets of the CCM-S. Dutifulness and Deliberation were assumed to represent both Control and Caution. Achievement Striving was taken to represent Industriousness. Finally, Self-Discipline on the NEO-PI-R was analogous to Procrastination Refrainment on the CCM-S.

Short almost perfect scale (SAPS)
The SAPS (Rice et al., 2014) is a short form of the Almost Perfect Scale Revised (APS-R; Slaney et al., 2001). It measures Perfectionism on two dimensions, namely Standards and Discrepancy, with four items per scale. The Standards dimension measures high performance expectations, whereas the Discrepancy dimension measures self-critical performance evaluations. English example items include "I expect the best from myself" (Standards) and "Doing my best never seems enough" (Discrepancy). The reported reliabilities are around $\alpha = .85$ for both subscales. The German- and French language-versions of the questionnaire were used.[4] All items were answered on a five-point Likert scale. Reliabilities were good for both scales (see Appendix). This

---

[3]Students in this sample were already assessed in grade 9 as part of the national ÉpStan cycle, and were hence included in our Study 1. However, only a subsample of these students could be matched longitudinally; we have therefore refrained from analyzing this selected subsample any further.

[4]The APS-R and SAPS are freely available in multiple languages from http://kennethwang.com/apsr/measures.html, last accessed 09.08.2020

Table 4. Correlations between the CCM-S and other personality questionnaires and reliabilities.

Scale	Subscale	Reliabilities		Correlations							
		α	ω	Industriousness	Caution	Control	Perfectionism	Procrastination Refrainment	Task Planning	Tidiness	Mean Conscientiousness
NEO-PI-R Conscientiousness	Competence	.61	.67	-.30***	-.45***	-.31***	-.39***	-.36***	-.39***	-.18***	-.51***
	Order	.63	.65	-.33***	-.36***	-.35***	-.29***	-.27***	**-.52***	**-.62***	-.62***
	Dutifulness	.58	.61	-.48***	**-.49***	**-.30***	-.44***	-.40***	-.44***	-.23***	-.60***
	Achievement Striving	.63	.69	**-.54***	-.39***	-.29***	-.58***	-.43***	-.58***	-.34***	-.69***
	Self-discipline	.81	.81	-.49***	-.41***	-.43***	-.40***	**-.69***	-.48***	-.35***	-.73***
	Deliberation	.72	.74	-.36***	**-.46***	**-.46***	-.32***	-.25***	-.44***	-.23***	-.54***
BFI-2-S	Extraversion	.68	.69	-.02	-.06	-.15**	-.19***	-.13***	-.15***	-.03	-.09
	Agreeableness	.48	.54	-.39***	-.39***	-.27***	-.04	-.25**	-.18**	-.11*	-.34***
	Conscientiousness	.69	.69	-.46***	-.43***	-.41***	-.36***	-.51***	-.51***	-.60***	-.74***
	Neuroticism	.50	.59	-.09	-.05	-.13**	-.16**	-.27***	-.04	-.10*	-.15**
	Openness to Experience	.35	.41	-.02	-.01	-.03	-.08	-.01	-.06	-.04	-.00
APS-R	Standards	.84	.85	-.47***	-.35***	-.11*	**-.61***	-.18***	-.50***	-.11*	-.49***
	Discrepancy	.83	.84	-.13**	-.02	-.19***	-.26***	-.15**	-.05	-.02	-.02

Note. NEO-PI-R = NEO-Personality Inventory-Revised Conscientiousness Scale, BFI-2-S = Big Five Inventory 2 short version, SAPS = Short Almost Perfect Scale, α = Cronbach's α, ω = McDonald's ω. Correlations with corresponding facet are indicated in bold. *p < .05, **p < .01, ***p < .001, no indication p > .05.

instrument was used to assess convergent validity for the Perfectionism facet of the CCM-S, which was considered analogous to the Standards facet of the APS-R.

GPA. Students self-reported their GPA on their most recent full-year report card, which was their Grade 9 GPA at the time of measurement. Grades in Luxembourg's secondary school system range from zero to 60, with higher grades indicating higher achievement.

Data analysis

Convergent validity was assessed through the correlations between the different CCM-S facets and corresponding scales of the NEO-PI-R or the SAPS (see Table 4). We further assessed the correlation between each facet and the overall Conscientiousness score of the BFI-2-S. Discriminant validity was assessed through the correlations between the facets and the non-Conscientiousness Big Five scales from the BFI-2-S. The correlations with Openness to Experience, Agreeableness, and Neuroticism are only reported for sake of completeness, as the reliabilities of the respective scales in our sample were too low for further use.

Relations with grades were assessed by calculating Pearson correlations between the mean scores on each individual facet and students' Grade 9 GPA. As we expected Industriousness to show a higher correlation with GPA than all other facets, we used the Fisher *r*-to-*z* transformation to assess significance of differences (Eid et al., 2015).

Results

All CCM-S facets showed significant positive associations with their corresponding scales in the NEO-PI-R or SAPS. All facets exhibited differential associations with the non-corresponding Conscientiousness facets. Caution and Control were not as clearly distinct as the other facets as they showed similar associations to the Deliberation and Dutifulness subscales of the NEO-PI-R. All facets exhibited significant positive associations with the Conscientiousness scale of the BFI-2-S, while showing descriptively smaller positive or non-significant correlations with the remaining Big Five constructs (see Table 4).

All facets except Tidiness were positively associated with Grade 9 GPA, with Industriousness, Perfectionism, and Caution exhibiting the relatively highest associations (Table 3). The correlation between Industriousness and GPA was significantly different from those to Task Planning (*p* = .026), Tidiness (*p* < .001), Procrastination Refrainment (*p* = .016), and Control (*p* = .019). Comparisons among the correlations of Caution, Perfectionism, and overall Conscientiousness with GPA revealed no significant differences (*p* > .05).

Discussion

The goal of Study 3 was to provide further evidence of validity of the CCM-S. Specifically, we assessed the convergent and discriminant validity of the different facets as well as their

criterion validity for GPA as an outcome measure. In line with our assumptions, all facets except Tidiness showed small to moderate positive associations with GPA. As expected and in line with previous research (MacCann et al., 2009; Rikoon et al., 2016), Industriousness had the descriptively strongest correlation with GPA out of all facets. Yet, Industriousness showed the same relation with GPA as broad Conscientiousness. Therefore, our results do not match the results from some studies that have found facets to be a better predictor of GPA than broad Conscientiousness (Paunonen & Ashton, 2001).

Overall, we were able to provide convincing evidence for the convergent and discriminant validity of the CCM-S. The associations between the individual facets and all discriminant Big Five traits closely resembled those of the original CCM (MacCann et al., 2009). The correlations between individual facets and broad Conscientiousness as measured by the BFI-2-S were moderate to high. While these numbers are smaller than the correlations reported for the original CCM (MacCann et al., 2009), it must be noted that the original study used the full version of the BFI (Benet-Martínez & John, 1998), whereas our study used the short version of the BFI-2. Examining the individual facets in more detail, the majority of analogous CCM-S and NEO-PI-R or APS-R facets showed moderate to high correlations. However, the Control facet showed only a small to moderate association with the Dutifulness facet of the NEO-PI-R. Although the Dutifulness facet arguably covers aspects of self-control that do not fully overlap with the Control facet of the CCM-S—i.e., the tendency to follow the rules vs. the tendency to control one's impulses—this result indicates that this particular facet scale covers a narrower concept of impulse control than other widely used instruments. Control might be more closely related to self-regulation than to Conscientiousness, causing the lack of convergence with potentially related constructs of Conscientiousness. For both self-control and self-regulation, impulse control seems to play a special role (see Inzlicht et al., 2021). Indeed, self-control in the CCM/CCM-S was split into self-reflection as measured by the Caution facet and in impulse control measured by the Control facet. This might lead to the finding that Caution is more closely related to Conscientiousness while Control is more closely related to self-regulation. Further research on the Control facet scale in isolation and its relation to Conscientiousness compared to self-regulation is therefore needed to clarify its role within the nomological network of Conscientiousness. Researchers planning to use this scale as a single measure of self-control should keep this in mind. In addition, as the Control facet also showed below acceptable model fit in Study 1, we recommend carefully assessing its psychometric properties when relying solely on this facet. Nonetheless, in light of the results for the overall scale, the Control facet can safely be used as a part of the overall CCM-S without compromising psychometric quality or content validity.

General discussion

In the present multi-study report, we developed and validated the CCM-S, a short measure for assessing seven facets of Conscientiousness in German and French. The CCM-S was developed specifically for research in educational settings and is suitable for use in large-scale assessments, as it consists of only four items for each facet. An extensive validation using student samples supported the psychometric quality of the overall scale and all facet scales. Only the Control facet needs further investigation before it can used as a standalone measurement instrument. We therefore recommend using the Control scale only as part of the overall CCM-S. There, it may provide incremental information regarding students' impulse control.

Our results can be seamlessly integrated into previous research on the facets of Conscientiousness. In line with this previous research, we found Industriousness to be the best predictors of GPA (MacCann et al., 2009; Rikoon et al., 2016). Likewise in accordance with previous results, we found Tidiness to have little to no predictive utility for academic achievement (MacCann et al., 2009). In addition, we found Caution to be the best predictor of SATS out of all CCM facets, which makes sense given its previously established link with intelligence (Rikoon et al., 2016). Given these findings, Conscientiousness, as a predictor of academic achievement, has the potential to support students in achieving their best in school. A solid understanding of the structure of Conscientiousness—that is, its lower-order facets—is necessary to convert this potential into actual practical usefulness. To better understand the role each facet plays in different academic outcomes, one could assess secondary school students' potential areas of weakness as well. This could, in turn, inform the development of interventions to foster specific facets of Conscientiousness among students to enhance academic outcomes. These trainings, if specifically targeted to compensate for students' weaknesses and reinforce their strengths, might support secondary school students in achieving their full potential in the long term (see e.g., Magidson et al., 2014). For this idea to be feasible in reality, further research on different facets of Conscientiousness in education is needed, especially in educational large-scale assessments. Initial results, including the present research, have shown that different facets of Conscientiousness exhibit differential utility for different academic outcomes (MacCann et al., 2009, 2015; Rikoon et al., 2016). However, studies on the individual facets are thus far mostly limited to cross-sectional investigations. Longitudinal panel studies, micro-longitudinal experience sampling, and intervention studies are required to uncover temporal, developmental, and causal aspects of the relationship between facets and outcomes. Moreover, the overall nomological network of individual Conscientiousness facets and the actual person-centered distribution of facets in students have not yet been investigated. The latter could reveal whether certain groups of students exhibit higher values on specific combinations of Conscientiousness facets, and how this translates into academic achievement. By presenting a concise, easy to implement instrument assessing seven facets of Conscientiousness, we hope to provide a foundation for future investigations.

We developed German and French versions of the CCM-S and established invariance across these language versions

for all facets. As the original CCM was developed based on English-language items from the IPIP (Goldberg et al., 2006), the CCM-S items are also available in English and can thus be used with English-speaking samples (see Appendix). However, it must be noted that we did not assess invariance between the English item versions and our translated items. Hence, the results presented for the German- and French-language versions of the CCM-S here might not be generalizable to the English items (Greiff & Scherer, 2018). Future studies should therefore address this issue by testing the invariance of the English wording used in the CCM compared to the German- and French-language versions of the CCM-S. Furthermore, an assessment of the CCM-S in different age groups would be valuable to test for potential age-specific variations.

We thus conclude that the CCM-S is a valid instrument that provides reliable scores. It was tailored specifically for use in (large-scale) educational assessments, measures seven facets of Conscientiousness with four items each and was developed based on fully representative student samples. Researchers may choose to include only one or more of the individual facets in their research, as we provided evidence for the psychometric quality for each separate scale. As noted earlier, only the Control facet scale should be treated cautiously when used individually. With this exception, both the overall instrument and the individual facet scales offer a solid foundation for future investigations of the lower-order facets of Conscientiousness.

Declaration of interest statement

We have no conflicts of interest to disclose.

Data availability statement

Our study uses data from the large-scale, Luxembourg School Monitoring Programme, the Épreuves Standardisées (ÉpStan). The ÉpStan has a proper legal basis and has been approved by the national committee for data protection. Appropriate ethical standards were followed in the conduct of the study (American Psychological Association, 2017). Participants and their parents/legal guardians were duly informed before the data collection and had the possibility to opt-out. All statistical analyses were performed with pseudonymized data, and a trusted-third-party-solution assured the privacy of the participants. Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data are not available.

References

Behr, D., Braun, M., & Dorer, B. (2016). Messinstrumente in internationalen Studien. GESIS Leibniz-Institut für Sozialwissenschaften (GESIS Survey Guidelines).

Benet-Martínez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the big five in Spanish and English. *Journal of Personality and Social Psychology*, 75(3), 729–750.

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, 130(6), 887–919.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, 37(4), 319–338. https://doi.org/10.1016/S0092-6566(02)00578-0

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Costa, P. T., & McCrae, R. R. (1992). Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. Psychological Assessment Ressources, Inc.

Danner, D., Rammstedt, B., Bluemke, M., Treiber, L., Berres, S., Soto, C. J., & John, O. P. (2016). Die deutsche version des big five inventory 2 (BFI-2). GESIS Leibniz-Institut für Sozialwissenschaften (GESIS Survey Guidelines).

de Vries, A., de Vries, R. E., & Born, M. P. (2011). Broad versus narrow traits: Conscientiousness and honesty-humility as predictors of academic criteria. *European Journal of Personality*, 25(5), 336–348. https://doi.org/10.1002/per.795

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence*, 15(2), 157–173. https://doi.org/10.1016/0160-2896(91)90028-C

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the big five. *Journal of Personality and Social Psychology*, 93(5), 880–896.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41(1), 417–440. https://doi.org/10.1146/annurev.ps.41.020190.002221

Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What no child left behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, 104(2), 439–451. https://doi.org/10.1037/a0026280

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *The Journal of Applied Psychology*, 91(1), 40–57.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology (London, England: 1953)*, 105(3), 399–412.

Eid, M., Gollwitzer, M., & Schmitt, M. (2015). Statistik und Forschungsmethoden: Mit Online-Materialien (4., überarb. und erw. Aufl.). Beltz.

Emons, W. H. M., Sijtsma, K., & Meijer, R. R. (2007). On the consistency of individual classification using short scales. *Psychological Methods*, 12(1), 105–120.

Enders, C. K. (2010). Applied missing data analysis. *Methodology in the social sciences*. Guilford Press.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, 15(6), 373–378. https://doi.org/10.1111/j.0956-7976.2004.00687.x

Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, 14(1), 47–64. https://doi.org/10.1016/j.lindif.2003.08.002

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Greiff, S., & Scherer, R. (2018). Still comparing apples with oranges? *European Journal of Psychological Assessment*, 34(3), 141–144. https://doi.org/10.1027/1015-5759/a000487

Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2007). Mechanisms by which childhood personality traits influence adult health status: Educational attainment and healthy behaviors. *Health Psychology*, 26(1), 121–125. https://doi.org/10.1037/0278-6133.26.1.121

Hayes, N., & Joseph, S. (2003). Big 5 correlates of three measures of subjective well-being. *Personality and Individual Differences*, 34(4), 723–727. https://doi.org/10.1016/S0191-8869(02)00057-0

Hu, Lt., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Inzlicht, M., Werner, K. M., Briskin, J. L., & Roberts, B. W. (2021). Integrating models of self-regulation. *Annual Review of Psychology*, 72, 319–345.

Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality*, 52, 29–36. https://doi.org/10.1016/j.jrp.2014.06.005

Jackson, J. J., Wood, D., Bogg, T., Walton, K. E., Harms, P. D., & Roberts, B. W. (2010). What do conscientious people do? Development and validation of the behavioral indicators of conscientiousness (BIC). *Journal of Research in Personality*, 44(4), 501–511.

Janssen, A. B., Schultze, M., & Grötsch, A. (2017). Following the ants. *European Journal of Psychological Assessment*, 33(6), 409–421. https://doi.org/10.1027/1015-5759/a000299

Kim, L. E., Poropat, A. E., & MacCann, C. (2016). Conscientiousness in Education: Its conceptualization, assessment, and utility. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *Psychosocial skills and school systems in the 21st century: Theory, research, and practice* (pp. 155–185). Springer series on human exceptionality. Springer.

Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence*, 36(2), 153–160. https://doi.org/10.1016/j.intell.2007.03.005

Kruyen, P. M., Emons, W. H. M., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing*, 13(3), 223–248. https://doi.org/10.1080/15305058.2012.703734

Leite, W. L., Huang, I.-C., & Marcoulides, G. A. (2008). Item selection for the development of short forms of scales using an ant colony optimization algorithm. *Multivariate Behavioral Research*, 43(3), 411–431.

Lounsbury, J. W., Steel, R. P., Loveland, J. M., & Gibson, L. W. (2004). An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence*, 33(5), 457–466. https://doi.org/10.1023/B:JOYO.0000037637.20329.97

MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences*, 19(4), 451–458. https://doi.org/10.1016/j.lindif.2009.03.007

MacCann, C., Lipnevich, A. A., Poropat, A. E., Wiemers, M. J., & Roberts, R. D. (2015). Self- and parent-rated facets of Conscientiousness predict academic outcomes: Parent-reports are more predictive, particularly for approach-oriented facets. *Learning*

and Individual Differences, 42, 19–26. https://doi.org/10.1016/j.lindif.2015.07.012

Magidson, J. F., Roberts, B. W., Collado-Rodriguez, A., & Lejuez, C. W. (2014). Theory-driven intervention for changing personality: Expectancy value theory, behavioral activation, and conscientiousness. *Developmental Psychology*, 50(5), 1442–1450.

Martin, R., Ugen, S., & Fischbach, A. (Eds.). (2014). Épreuves Standardisées. Bildungsmonitoring für Luxemburg. Nationaler Bericht 2011 | 2013 [Épreuves Standardisées: School monitoring for Luxembourg. National report 2011 to 2013]. Men.lu—Website of the Ministry of National Education, Children and Youth http://www.men.public.lu/catalogue-publications/secondaire/statistiques-analyses/autres-themes/epreuves-standard-11-13/epstan.pdf

McDonald, R. P. (1999). *Test theory: A unified treatment*. Erlbaum.

Meldrum, R. C., Petkovsek, M. A., Boutwell, B. B., & Young, J. T. N. (2017). Reassessing the relationship between general intelligence and self-control in childhood. *Intelligence*, 60, 1–9. https://doi.org/10.1016/j.intell.2016.10.005

Migali, G., & Zucchelli, E. (2017). Personality traits, forgone health care and high school dropout: Evidence from US adolescents. *Journal of Economic Psychology*, 62, 98–119. https://doi.org/10.1016/j.joep.2017.06.007

Muthén, L. K., & Muthén, B. O. (1998/2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Niepel, C., Greiff, S., Mohr, J. J., Fischer, J.-A., & Kranz, D. (2019). The English and German versions of the Lesbian, Gay, and Bisexual Identity Scale: Establishing measurement invariance across nationality and gender groups. *Psychology of Sexual Orientation and Gender Diversity*, 6(2), 160–174. https://doi.org/10.1037/sgd0000315

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, 93(1), 116–130. https://doi.org/10.1037/0022-3514.93.1.116

Olaru, G., Witthöft, M., & Wilhelm, O. (2015). Methods matter: Testing competing models for designing short-scale big-five assessments. *Journal of Research in Personality*, 59, 56–68. https://doi.org/10.1016/j.jrp.2015.09.001

Organization for Economic Co-operation and Development. (2016). *Global competency for an inclusive world*. OECD Publishing. https://www.oecd.org/education/Global-competency-for-an-inclusive-world.pdf

Paunonen, S. V., & Ashton, M. C. (2001). Big five predictors of academic achievement. *Journal of Research in Personality*, 35(1), 78–90. https://doi.org/10.1006/jrpe.2000.2309

Paunonen, S. V., & Ashton, M. C. (2013). On the prediction of academic performance with personality traits: A replication study. *Journal of Research in Personality*, 47(6), 778–781. https://doi.org/10.1016/j.jrp.2013.08.003

Peabody, D., & de Raad, B. (2002). The substantive nature of psycholexical personality factors: A comparison across languages. *Journal of Personality and Social Psychology*, 83(4), 983–997. https://doi.org/10.1037/0022-3514.83.4.983

Perugini, M., & Gallucci, M. (1997). A hierarchical faceted model of the big five. *European Journal of Personality*, 11(4), 279–301. https://doi.org/10.1002/(SICI)1099-0984(199711)11:4 < 279::AID-PER282 > 3.0.CO;2-F

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338. https://doi.org/10.1037/a0014996

Poropat, A. E. (2014a). A meta-analysis of adult-rated child personality and academic performance in primary education. *The British Journal of Educational Psychology*, 84(Pt 2), 239–252. https://doi.org/10.1111/bjep.12019

Poropat, A. E. (2014b). Other-rated personality and academic performance: Evidence and implications. *Learning and Individual Differences*, 34, 24–32. https://doi.org/10.1016/j.lindif.2014.05.013

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review: DR*, 41, 71–90.

R Core Team. (2007). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, *25*(4), 1137–1145.

Rice, K. G., Richardson, C. M. E., & Tueller, S. (2014). The short form of the revised almost perfect scale. *Journal of Personality Assessment*, *96*(3), 368–379.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387.

Rikoon, S. H., Brenneman, M., Kim, L. E., Khorramdel, L., MacCann, C., Burrus, J., & Roberts, R. D. (2016). Facets of conscientiousness and their differential relationships with cognitive ability factors. *Journal of Research in Personality*, *61*, 22–34. https://doi.org/10.1016/j.jrp.2016.01.002

Roberts, B. W., Bogg, T., Walton, K. E., Chernyshenko, O. S., & Stark, S. E. (2004). A lexical investigation of the lower-order structure of conscientiousness. *Journal of Research in Personality*, *38*(2), 164–178. https://doi.org/10.1016/S0092-6566(03)00065-5

Roberts, B. W., Chernyshenko, O. S., Stark, S. E., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, *58*(1), 103–139. https://doi.org/10.1111/j.1744-6570.2005.00301.x

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In *Handbook of individual differences in social behavior* (pp. 369–381). The Guilford Press.

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the big five personality factors: A cross-language replication. *Journal of Personality and Social Psychology*, *76*(4), 613–627. https://doi.org/10.1037/0022-3514.76.4.613

Schroeders, U., Wilhelm, O., & Olaru, G. (2016). The influence of item sampling on sex differences in knowledge tests. *Intelligence*, *58*, 22–32. https://doi.org/10.1016/j.intell.2016.06.003

Slaney, R. B., Rice, K. G., Mobley, M., Trippi, J., & Ashby, J. S. (2001). The revised almost perfect scale. *Measurement and Evaluation in Counseling and Development*, *34*(3), 130–145. https://doi.org/10.1080/07481756.2002.12069030

Smith, J., Ryan, L. H., & Röcke, C. (2013). The day-to-day effects of conscientiousness on well-being. *Research in Human Development*, *10*(1), 9–25. https://doi.org/10.1080/15427609.2013.760257

Song, J., Gaspard, H., Nagengast, B., & Trautwein, U. (2020). The Conscientiousness × interest compensation (CONIC) model: Generalizability across domains, outcomes, and predictors. *Journal of Educational Psychology*, *112*(2), 271–287. https://doi.org/10.1037/edu0000379

Soto, C. J., & John, O. P. (2017a). The next big five inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Soto, C. J., & John, O. P. (2017b). Short and extra-short forms of the big five inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69–81. https://doi.org/10.1016/j.jrp.2017.02.004

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*(4), 718–737.

Takahashi, Y., Edmonds, G. W., Jackson, J. J., & Roberts, B. W. (2013). Longitudinal correlated changes in conscientiousness, preventative health-related behaviors, and self-perceived physical health. *Journal of Personality*, *81*(4), 417–427. https://doi.org/10.1111/jopy.12007

Trautwein, U., Lüdtke, O., Nagy, N., Lenski, A., Niggli, A., & Schnyder, I. (2015). Using individual interest and conscientiousness to predict academic effort: Additive, synergistic, or compensatory effects? *Journal of Personality and Social Psychology*, *109*(1), 142–162. https://doi.org/10.1037/pspp0000034

Yarkoni, T. (2010). The abbreviation of personality, or how to measure 200 personality scales with 200 items. *Journal of Research in Personality*, *44*(2), 180–198. https://doi.org/10.1016/j.jrp.2010.01.002

Appendix

CCM-S

German and French versions of the CCM-S, as well as corresponding English wordings from the CCM (MacCann et al., 2009), which was developed using items from the IPIP (Goldberg et al., 2006).

Item number	Facet	English wording	German version	French version	Reverse coded
1	Industriousness	I am always prepared.	Ich bin immer vorbereitet.	Je suis toujours préparé(e).	NO
2	Industriousness	I do more than what's expected of me.	Ich mache mehr als von mir erwartet wird.	J'en fais plus que ce que l'on attend de moi.	NO
3	Industriousness	I make an effort.	Ich bemühe mich.	Je fais un effort.	NO
4	Industriousness	I work hard.	Ich arbeite viel.	Je travaille dur.	NO
5	Caution	I behave properly.	Ich verhalte mich angemessen.	Je me comporte convenablement.	NO
6	Caution	I look at the facts.	Ich berücksichtige die Fakten.	Je regarde les faits.	NO
7	Caution	I make careful choices.	Ich treffe sorgfältige Entscheidungen.	Je fais des choix prudents.	NO
8	Caution	I think ahead.	Ich denke vorausschauend.	J'anticipe.	NO
9	Control	I act impulsively when something is bothering me.	Ich handle impulsiv, wenn mich etwas stört.	J'agis avec impulsivité quand quelque chose me dérange.	YES
10	Control	I do unexpected things.	Ich mache unerwartete Dinge.	Je fais des choses inattendues.	YES
11	Control	I make a fool of myself.	Ich mache mich lächerlich.	Je me ridiculise.	YES
12	Control	I make rash decisions.	Ich treffe unüberlegte Entscheidungen.	Je prends des décisions sans réfléchir.	YES
13	Perfectionism	I continue until everything is perfect.	Ich bleibe so lange an etwas dran, bis alles perfekt ist.	Je continue jusqu'à ce que tout soit parfait.	NO
14	Perfectionism	I detect mistakes.	Ich entdecke Fehler.	Je trouve les erreurs.	NO
15	Perfectionism	I go straight for the goal.	Ich gehe geradewegs aufs Ziel zu.	Je vais droit au but.	NO
16	Perfectionism	I try to outdo others.	Ich versuche, andere zu übertreffen.	J'essaie de faire mieux que les autres.	NO
17	Procrastination Refrainment	I am easily distracted.	Ich lasse mich leicht ablenken.	Je me laisse facilement distraire.	YES
18	Procrastination Refrainment	I have difficulty starting tasks.	Es fällt mir schwer, mit Aufgaben zu beginnen.	J'ai du mal à commencer des tâches.	YES
19	Procrastination Refrainment	I put off unpleasant tasks.	Ich schiebe unangenehme Aufgaben auf.	Je remets les tâches désagréables à plus tard.	YES
20	Procrastination Refrainment	I waste my time.	Ich verschwende meine Zeit.	Je perds mon temps.	YES
21	Task planning	I am a goal-oriented person.	Ich bin ein zielstrebiger Mensch.	Je suis une personne déterminée.	NO
22	Task planning	I do things according to a plan.	Ich erledige Dinge nach einem Plan.	Je fais les choses comme c'était planifié.	NO
23	Task planning	I like to plan ahead.	Ich plane gerne voraus.	J'aime planifier à l'avance.	NO
24	Task planning	I make plans and stick to them.	Ich mache Pläne und halte mich daran.	Je fais des plans et je m'y tiens.	NO
25	Tidiness	I am not bothered by messy people.	Unordentliche Menschen stören mich nicht.	Les gens désordonnés ne me dérangent pas.	YES
26	Tidiness	I leave a mess in my room.	In meinem Zimmer herrscht großes Chaos.	Ma chambre est en grande désordre.	YES
27	Tidiness	I leave my belongings around.	Ich lasse meine Sachen herumliegen.	Je laisse traîner mes affaires.	YES
28	Tidiness	I often forget to put things in their proper place.	Ich vergesse oft, die Dinge wieder an ihren Platz zurück zu legen.	J'oublie souvent de ranger les choses à leur place.	YES

Note. All items are available for free and open source.