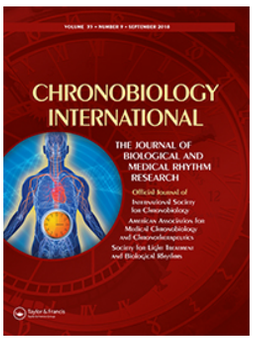
Convergent and construct validity and test–retest reliability of the Caen Chronotype Questionnaire in six languages



Taylor & Francis
Taylor & Francis Group

Chronobiology International

The Journal of Biological and Medical Rhythm Research

ISSN: 0742-0528 (Print) 1525-6073 (Online) Journal homepage: http://www.tandfonline.com/loi/icbi20

Convergent and construct validity and test–retest reliability of the Caen Chronotype Questionnaire in six languages

Sylvain Laborde, Fabrice Dosseville, Asma Aloui, Helmi BEN SAAD, Maurizio Bertollo, Laura Bortoli, Barbara Braun, Karim Chamari, Hamdi Chtourou, Yvonne De Kort, Abdulaziz Farooq, Marijke CM Gordijn, Pablo Greco, Félix Guillén, Monoem Haddad, Thomas Hosang, Karim Khalladi, Romain Lericollais, Mariana Lopes, Claudio Robazza, Karin Smolders, Alexander Wurm & Mark S Allen

To cite this article: Sylvain Laborde, Fabrice Dosseville, Asma Aloui, Helmi BEN SAAD, Maurizio Bertollo, Laura Bortoli, Barbara Braun, Karim Chamari, Hamdi Chtourou, Yvonne De Kort, Abdulaziz Farooq, Marijke CM Gordijn, Pablo Greco, Félix Guillén, Monoem Haddad, Thomas Hosang, Karim Khalladi, Romain Lericollais, Mariana Lopes, Claudio Robazza, Karin Smolders, Alexander Wurm & Mark S Allen (2018) Convergent and construct validity and test–retest reliability of the Caen Chronotype Questionnaire in six languages, Chronobiology International, 35:9, 1294-1304, DOI: 10.1080/07420528.2018.1475396

To link to this article: https://doi.org/10.1080/07420528.2018.1475396

CHRONOBIOLOGY INTERNATIONAL
2018, VOL. 35, NO. 9, 1294–1304
https://doi.org/10.1080/07420528.2018.1475396

Taylor & Francis
Taylor & Francis Group

Convergent and construct validity and test–retest reliability of the Caen Chronotype Questionnaire in six languages

Sylvain Laborde [a,b], Fabrice Dosseville[b], Asma Aloui[c], Helmi BEN SAAD [d], Maurizio Bertollo[e], Laura Bortoli[e], Barbara Braun[a], Karim Chamari [f], Hamdi Chtourou[g], Yvonne De Kort[h], Abdulaziz Farooq[f], Marijke CM Gordijn[i], Pablo Greco[j], Félix Guillén[k], Monoem Haddad [l], Thomas Hosang[m], Karim Khalladi[f], Romain Lericollais[b], Mariana Lopes[n], Claudio Robazza [e], Karin Smolders[h], Alexander Wurm[a], and Mark S Allen[o]

[a]Institute of Psychology, German Sport University, Cologne, Germany; [b]EA 4260, UFR STAPS, Université de Caen Basse-Normandie, Caen, France; [c]High Institute of Sport and Physical Education, Universite of Gafsa, Gafsa, Tunisia; [d]Heart Failure Research Laboratory (LR12SP09), Farhat HACHED Hospital, Universite de Sousse, Sousse, Tunisia; [e]Department of Medicine and Aging Sciences, BIND-Behavioral Imaging and Neural Dynamics Center, Chieti, Italy; [f]Aspetar, Qatar Orthopaedic and Sports Medicine Hospital, Doha, Qatar; [g]High Institute of Sport and Physical Education of Sfax, Universite de Sfax, Sfax, Tunisia; [h]Human Technology Interaction, Eindhoven University of Technology, Eindhoven, Netherlands; [i]Chrono@Work B.V., Groningen, Netherlands; [j]Center for studies in Cognition and Action, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; [k]University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain; [l]Sport Science Program, College of Arts and Sciences, Qatar University, Doha, Qatar; [m]Deutsche Sporthochschule Koln, Psychologisches Institut, Koln, Germany; [n]Physical Education Department of Universidade Federal de Viçosa, Research Group in Action and Cognition, Viçosa, Brazil; [o]School of Psychology, University of Wollongong, Wollongong, Australia

ABSTRACT

Chronotype questionnaires provide a simple and time-effective approach to assessing individual differences in circadian variations. Chronotype questionnaires traditionally focused on one dimension of chronotype, namely its orientation along a continuum of morningness and eveningness. The Caen Chronotype Questionnaire (CCQ) was developed to assess an additional dimension of chronotype that captures the extent to which individual functioning varies during the day (amplitude). The aim of this study was to provide a multilanguage validation of the CCQ in six world regions (Arabic, Dutch, German, Italian, Portuguese and Spanish). At Time 1, a total of 2788 participants agreed to take part in the study (Arabic, n = 731; Dutch, n = 538; German, n = 329; Italian, n = 473; Portuguese, n = 361; Spanish, n = 356). Participants completed an assessment of the CCQ together with the Morningness-Eveningness Questionnaire (MEQ; Horne & Ostberg 1976) as well as questions related to factors theoretically related to chronotype (age, shift work, physical activity, sleep parameters and coffee consumption). One month later, participants again completed the CCQ. Results showed that the two-factor structure (morningness-eveningness and amplitude) of the CCQ could be replicated in all six languages. However, measurement invariance could not be assumed regarding the factor loadings across languages, meaning that items loaded more on their factors in some translations than in others. Test–retest reliability of the CCQ ranged from unacceptable (German version) to excellent (Dutch, Portuguese). Convergent validity was established through small–medium effect size correlations between the morningness-eveningness dimension of the CCQ and the MEQ. Taken together, our findings generally support the use of the translated versions of the CCQ. Further validation work on the CCQ is required including convergent validation against physiological markers of sleep, health and well-being.

ARTICLE HISTORY
Received 19 March 2018
Revised 7 May 2018
Accepted 9 May 2018

KEYWORDS
Confirmatory factor analysis; sleep; questionnaire validation; adaptation; personality

Introduction

If recognizing that some people function better as either an early bird or night owl is rather easy, knowing how strongly they function in this manner might not be so obvious. These two aspects represent two dimensions of chronotype (that represent individual differences in circadian variations) in terms of morningness-eveningness and amplitude (Horne and Ostberg 1976; Dosseville et al. 2013). The morningness-eveningness dimension reflects better individual functioning in the morning or in the evening, whereas the amplitude dimension reflects how strong the difference in functioning during the daytime actually is. The Caen Chronotype Questionnaire (CCQ; Dosseville et al. 2013) was developed to assess these two dimensions of chronotype for use in the French language. Initial validation work suggested

CONTACT Sylvain Laborde s.laborde@dshs-koeln.de
Supplemental data for this article can be accessed here

that the questionnaire was a valid tool for accurately measuring the two dimensions of chronotype. However, the test–retest reliability of the measures has not been established, and the validity of the measures in other languages is unknown. The aim of this study was to translate the questionnaire into several languages and to establish test–retest reliability and convergent and construct validity of the measures across these languages.

In the circadian rhythm, the morningness-eveningness dimension reflects the phase aspect, which represents the peak time of the circadian variation (Dosseville et al. 2013; Refinetti et al. 2007). Morningness-eveningness has a strong biological basis (Von Schantz et al. 2015) and is related to aspects of human personality and health (Adan et al. 2012). Morningness refers to people who wake early in the morning, who are in a better mood and work more efficiently in the morning hours (Ogińska 2011). Overall, these individuals feel more energetic in the early hours of the day and usually go to bed early at night. Morning people usually show a higher degree of conscientiousness (Tsaousis 2010) as well as more health-conscious behavior (Merikanto et al. 2013). Eveningness refers to people who perform better later in the day (Ogińska 2011). So-called night owls rarely wake early and tend to go to bed late. Evening people tend to show higher levels of extraversion, impulsivity, novelty seeking and open-mindedness (Adan et al. 2012). Eveningness is also related to a greater likelihood of physical and mental health problems (Partonen 2015) including, for example, depression (Randler et al. 2012) and negative affect (Simor et al. 2015).

In the circadian rhythm, amplitude corresponds to the half of the peak-to-trough variation (Refinetti et al. 2007). In other words, amplitude is "the range of diurnal fluctuations in morningness–eveningness" (Randler et al. 2016). The amplitude dimension reflects the strength of the human circadian system, this is to say, how strong the changes in performance and mood are during the day (Aschoff and Pohl 1978; Dosseville et al. 2013). As assessed by questionnaire, the amplitude dimension allows capturing the subjective feeling of the magnitude of diurnal variations (Oginska et al. 2017). The amplitude dimension can help predict the height of performance peaks and also hypo- and hyperactivation phases that can be used for modulating psychophysical states (Ogińska 2011). For example, knowing better one's chronotype amplitude may help to optimize the ability to relax in stressful situations or to activate oneself during phases of drowsiness during the day. Research on the amplitude dimension has found, for example, that greater diurnal fluctuations (higher amplitude) are associated with a greater risk of mood and anxiety disorders (Nowakowska-Domagala et al. 2016) and lower participation rates in sport and physical activity (Laborde et al. 2015).

For a long time research on chronotype focused exclusively on the morningness-eveningness dimension, and the investigation of the amplitude dimension via questionnaire first appeared in the Chronotype Questionnaire (Ogińska 2011). This questionnaire aimed to assess two dimensions: the morningness-eveningness dimension and a distinctness dimension (this latter term was used by Oginska to depict the amplitude dimension). In an attempt to translate and validate the Chronotype Questionnaire to the French language, Dosseville and collaborators (2013) showed a poor fit to the factor structure in a sample of younger and older adults. This led to a revision of the Chronotype Questionnaire that involved shortening the questionnaire to 18 items measuring the morningness-eveningness dimension, and modifications to the distinctness dimension—subsequently referred to as amplitude. In their second study (Dosseville et al. 2013), factor analysis was used to shorten the questionnaire to 16 items and two dimensions, morningness-eveningness and amplitude. These two factors were confirmed in a third study using confirmatory factor analysis, and convergent validity of the morningness-eveningness dimension was established using the Morningness-Eveningness Questionnaire (MEQ; Horne and Ostberg 1976). The new 16-item questionnaire was named the CCQ (Dosseville et al. 2013).

So far, the CCQ is only available in the French (Dosseville et al. 2013) and Polish (Ogińska et al. 2017) languages. A cross-cultural study used a German, Indian and Slovakian version of the CCQ (Randler et al. 2015), but no information was provided regarding the validation process of the CCQ into these languages (e.g. translation procedure, confirmatory factor analysis and test–retest reliability). Following guidelines regarding test adaptations, describing

those steps is crucial to be able to interpret correctly and to have confidence in the findings obtained with the adapted scales (Hambleton & De Jong 2003). Therefore, work is still needed to provide a complete validation process and verify the two-factor structure of the CCQ in different languages. The aim of this study was to extend the potential use of the CCQ by validating it in several languages (viz., Arabic, Dutch, German, Italian, Portuguese and Spanish). Combined, these languages represent more than 1.1 billion first-language speakers (Ethnologue 2018). We first tested the two-factor structure across the six translations using confirmatory factor analysis. We also tested the reliability of the CCQ using a 1-month test–retest approach. This time period was chosen given the relative stability of human chronotype over short timespans (Horne and Ostberg 1976; Dosseville et al. 2013). We further tested the convergent validity of the CCQ against the most widely used measure of chronotype—the MEQ. Similar to the original validation study (Dosseville et al. 2013), we anticipated high correlations between the CCQ-ME and the MEQ, but not between the CCQ-A and the MEQ. In addition, to further establish criterion validity of the measure, several known correlates of chronotype were also assessed: shift work (e.g. Wickwire et al. 2017), sleep duration (e.g. Malone et al. 2017), napping (e.g. Lee et al. 2017; Suh et al. 2017), sleep treatment (e.g. Coogan and McGowan 2017) and coffee consumption (e.g. Suh et al. 2017). Small significant correlations between chronotype and these parameters would be considered further evidence for the criterion validity of the questionnaire.

Material and methods

Participants

A total of 2788 participants took part in the study and sample characteristics are displayed in Table 1. The participants were from seven countries: Qatar and Tunisia for Arabic, the Netherlands for Dutch, Germany for German, Italy for Italian, Brazil for Portuguese and Spain for Spanish. Data collection occurred between 2014 and 2017. Ethical approval was obtained prior to data collection via the university research ethics committee of the first author, and the protocol complied with the ethical standards of the journal for the conduct of human biological rhythm research (Portaluppi et al. 2010).

Questionnaires

The Caen Chronotype Questionnaire

The CCQ (Dosseville et al. 2013) consists of 16 items measuring two dimensions of chronotype. The two dimensions are morningness-eveningness (CCQ-ME, eight items) and amplitude (CCQ-A, eight items). Both dimensions are measured on a scale from 1 (*totally disagree*) to 5 (*totally agree*). Item examples include: "I feel I can think the best in the morning" (CCQ-ME) and "I can work efficiently at any time of the day" (CCQ-A).

The Morningness-Eveningness Questionnaire

The MEQ (Horne and Ostberg 1976) measures only the morningness-eveningness dimension of chronotype, that refers to the phase of an individual's sleep-wake cycle, and is conceptually similar to the CCQ-ME (Dosseville et al. 2013). The MEQ consists of 19 items answered either by single choice or by pointing a certain time. The answers are coded from 1 to 4, 1 to 5, 1 to 6, 0 to 5, or 0 to 6 according to the question. An item example is: "at what time would you get up if you were entirely free to plan your day?"

Sociodemographic and lifestyle factors

Participants were asked about their physical activity involvement ("How many hours do you exercise per week?") and usual sleep duration ("How many hours do you usually sleep per night?"), and were required to provide a numerical answer in hours. They were also asked about their involvement in shift work ("If you have a job, is it based on shift work?"), their tendency to take naps during the day ("Do you usually take a nap during the day?"), and whether they were taking any medical treatment for sleep ("Do you have a treatment for sleep?"), with all answers given as either "yes" or "no". Participants were also asked to provide the average number of coffee cups they drink during the day, with a numerical answer provided. These variables were chosen based on past research that has reported small positive correlations between these variables and dimensions of chronotype (see e.g. Mota et al. 2016; Whitter et al. 2014).

Table 1. Sample demographic characteristics.

		Arabic		Dutch		German		Italian		Portuguese		Spanish	
Time 1	N	731		538		329		473		361		356	
	Male	418		188		222		242		179		183	
	Female	313		350		107		231		182		173	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
	CCQ-ME	2.6	0.8	3.1	0.7	3.1	0.7	2.5	0.8	3.1	0.8	2.9	0.7
	CCQ-A	3.2	0.7	3.0	0.6	3.0	0.6	2.9	0.7	3.0	0.8	3.0	0.6
	MEQ	–	–	50.7	3.9	50.0	7.8	50.4	4.0	–	–	51.7	9.1
	Age	25.1	8.1	46.3	15.5	21.7	3.8	43.3	14.3	24.0	7.1	29.6	10.5
	Age range	18–67		18–79		18–60		19–75		18–56		18–60	
	Sleep quantity (hours per night)	7.4	1.4	7.7	3.3	7.5	0.8	7.1	1.1	7.0	1.2	7.3	1.1
	Coffee cups (per day)	1.5	1.2	3.0	2.7	0.8	1.3	2.4	1.6	1.5	1.7	1.0	1.2
	Physical activity (hours per week)	2.8	4.1	4.1	4.8	10.9	5.9	1.9	3.2	4.1	4.9	3.9	4.1
		N	%	N	%	N	%	N	%	N	%	N	%
	Shift work	371	50.8	55	10.2	3	0.9	107	22.6	121	33.5	65	18.3
	Nap	222	30.3	95	17.7	31	9.4	121	25.6	130	36.0	144	40.5
	Sleep medicine	336	46.0	19	3.5	0	0.0	11	2.3	16	4.4	44	12.4
Time 2	N	506		208		161		284		157		174	
	Male	307		75		103		150		82		90	
	Female	199		133		58		134		75		84	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
	CCQ-ME	2.8	0.9	3.0	0.6	3.0	0.6	2.6	0.8	3.1	0.9	3.0	0.7
	CCQ-A	3.2	0.7	3.4	0.5	3.4	0.5	2.9	0.6	3.1	0.8	3.0	0.6
N Ratio	%	69.2		38.7		48.9		60.0		43.5		48.9	

MEQ: Morningness-Eveningness Questionnaire; CCQ: Caen Chronotype Questionnaire; CCQ-ME: Morningness-eveningness scale of the CCQ; CCQ-A: Amplitude scale of the CCQ; Time 1 corresponds to filling out the demographic questions, the Caen Chronotype Questionnaire, and the Morningness-Eveningness Questionnaire; while Time 2 corresponds to filling out the Caen Chronotype Questionnaire 1 month later; *N* Ratio: percentage of participants who responded both at Times 1 and 2.

Procedure

Following guidelines in test adaptation (Hambleton and De Jong 2003), the CCQ was first translated into six languages (Arabic, Dutch, German, Italian, Portuguese and Spanish). The basis for the CCQ translations was the English version of the CCQ (Dosseville et al. 2013).[1] The CCQ was first translated by two independent bilingual persons for each of the languages, native speakers of the language in which the questionnaire was being translated, and perfectly fluent in English. Two other translators with the same characteristics translated it back to English again. Item wording was then discussed with the first two translators until consensus was reached. Data collection was performed with both paper-pencil and online surveys for all languages. These two modes of data collection provide similar results (Seo and De Jong 2015). For each language, all questionnaires were completed at Time 1 and the CCQ was completed again 1 month later (Time 2).

Data collection at Time 1 took approximately 20 min and data collection at Time 2 took approximately 5 min. Participants received no compensation for their voluntary participation, except for the Dutch sample (i.e. monetary reward). The MEQ was included for all languages except Arabic and Portuguese, given that we could not locate a MEQ version in these two languages. Given that the holy month of Ramadan is accompanied with substantial changes in lifestyle, including sleeping pattern in (i) fasting Muslims and even in (ii) nonfasting individuals living in Muslim-majority countries, the study was not conducted during Ramadan or the month following Ramadan. Data collection was anonymous. A

[1]We should note at this stage that the English version provided in the original validation study of the CCQ did not undergo a full validation procedure, only the translation and back-translation steps. The current investigation also intended to include an English language sample. Data were collected from 365 Australian adults at Time 1 with 337 completing follow-up questionnaires 1 month later. However, due to human error, one question from the CCQ was not included (at either time point) and, therefore, it was necessary to discard this data for the purpose of validation.

code was used to match the data for the CCQ from Time 1 to Time 2. Participants could cease their participation in the study at any time without giving any reason.

Data analysis

The data were first checked for normality and outliers. To investigate whether the two-factor structure could be confirmed in each language, we ran separate confirmatory factor analyses using IBM AMOS 25.0 (New York, USA). To test for measurement invariance across versions, we then ran a multigroup comparison confirmatory factor analysis. For the separate confirmatory factor analyses, we used the following indices to evaluate goodness-of-fit: the χ^2 statistic, the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR). The χ^2 statistic provides a subjective index of fit, with large χ^2 values relative to degrees of freedom indicating a poor fit, and small values indicating a good fit (Jöreskog 1993). For the CFI and the TLI, values between 0.90 and 0.94 indicate acceptable fit and values of 0.95 and higher indicate a good fit (Hu and Bentler 1999). For the RMSEA, values below 0.08 indicate acceptable fit, whereas values below 0.05 indicate good fit. For the SRMR, a good fit is indicated by values smaller than 0.08 (Hu and Bentler 1999).

For the multigroup analysis, to test the difference between two invariance models, we do not rely on the χ^2 difference, as it is too restrictive, and instead we used the change in CFI value (Byrne 2009), which has to be lower than 0.01 to demonstrate invariance between two models (Cheung and Rensvold 2002). Test–retest reliability, convergent validity with the MEQ, and the relationships with sociodemographic and lifestyle factors were investigated using bivariate correlations. For criterion validity, a correlation < 0.10 is considered trivial, 0.11–0.29 is considered small, 0.30–0.49 is considered medium and ≥ 0.50 is considered large (Cohen 1988). For test–retest reliability, a correlation of 1.00 indicates perfect reliability, a correlation of ≥ 0.90 indicates excellent reliability, a correlation of ≥ 0.80, < 0.90 indicates good reliability, a correlation of ≥ 0.70, < 0.80 indicates acceptable reliability, a correlation of ≥ 0.60, < 0.70 indicates questionable reliability, a correlation of ≥ 0.50, < 0.60 indicates poor reliability and a correlation of < 0.50 indicates unacceptable reliability. Finally, concerning the CCQ-ME and CCQ-A, the cutoff values for each quartile are presented in Table 2.

Results

Confirmatory factor analyses

Our participant-per-item ratio for each version was above the minimum 10:1 ratio recommended (e.g. Nunally 1978). Table 3 shows that for all samples there was a good fit to the theoretically expected two-factor structure. Taken individually, the data within each adapted version fit the two-

Table 2. Quartile distribution for the morningness-eveningness and amplitude dimensions of the Caen Chronotype Questionnaire.

		Arabic	Dutch	German	Italian	Portuguese	Spanish
CCQ morningness-eveningness	25%	2.00	2.13	2.63	2.00	2.50	2.38
	50%	2.50	2.75	3.00	2.38	3.00	2.88
	75%	3.13	3.50	3.50	3.00	3.75	3.50
CCQ amplitude	25%	2.75	2.50	2.50	2.50	2.50	2.63
	50%	3.25	3.13	3.00	2.88	3.00	3.00
	75%	3.75	3.63	3.38	3.25	3.63	3.38

CCQ: Caen Chronotype Questionnaire.

Table 3. Separate confirmatory factor analyses across languages.

	χ^2	DF	*p*	TLI	CFI	RMSEA (90% CI)	SRMR
Arabic	140.865	103	0.008	0.97	0.98	0.022 (0.015; 0.073)	0.04
Dutch	163.626	103	< 0.001	0.97	0.97	0.040 (0.028; 0.051)	0.06
German	132.692	103	0.026	0.96	0.96	0.030 (0.011; 0.043)	0.05
Italian	160.158	103	< 0.001	0.94	0.95	0.034 (0.023; 0.044)	0.05
Portuguese	146.739	103	0.003	0.95	0.96	0.034 (0.021; 0.046)	0.06
Spanish	133.955	103	0.022	0.96	0.97	0.029 (0.012; 0.042)	0.05

TLI: Tucker–Lewis index; CFI: Comparative fit index; RMSEA: Root mean square of approximation; SRMR: Standardized root mean square residual.

Table 4. Factor loadings.

	Item	Arabic	Dutch	German	Italian	Portuguese	Spanish
Amplitude	Item 1	0.54	0.66	0.47	0.54	0.59	0.52
	Item 3	0.35	0.46	0.35	0.34	0.38	0.32
	Item 6	0.44	0.57	0.37	0.46	0.48	0.44
	Item 8	0.64	0.83	0.57	0.66	0.71	0.60
	Item 10	0.48	0.67	0.46	0.54	0.55	0.51
	Item 12	0.40	0.56	0.36	0.43	0.46	0.46
	Item 14	0.50	0.63	0.47	0.52	0.51	0.47
	Item 15	0.50	0.69	0.55	0.52	0.56	0.49
Morningness-Eveningness	Item 2	0.63	0.74	0.57	0.60	0.66	0.68
	Item 4	0.55	0.64	0.62	0.53	0.52	0.58
	Item 5	0.49	0.67	0.48	0.50	0.47	0.55
	Item 7	0.41	0.51	0.34	0.45	0.43	0.44
	Item 9	0.55	0.72	0.57	0.53	0.58	0.56
	Item 11	0.40	0.68	0.43	0.55	0.55	0.60
	Item 13	0.45	0.59	0.39	0.46	0.53	0.47
	Item 16	0.56	0.75	0.69	0.67	0.67	0.65

factor structure of the CCQ. We wanted to test whether measurement invariance could be assumed among the different versions. The unconstrained model of the multigroup analysis, pooling all data together, showed a good fit to the data and confirmed that the data for the whole sample (n = 2788) fit the two-factor structure: χ^2 = 876.136, DF = 618, $p < 0.001$, CFI = 0.97, TLI = 0.96, RMSEA = 0.013 (90% CI: 0.011; 0.015), SRMR = 0.04. However, the CFI difference between the unconstrained model and the measurement weights model (0.06), was higher than 0.01, χ^2 = 1360.818, DF = 688, $p < 0.001$, CFI = 0.91, TLI = 0.91, RMSEA = 0.019 (90% CI: 0.018; 0.021), SRMR = 0.05. This means that measurement invariance cannot be assumed at the level of the factor loadings. Factor loadings across the six samples are provided in Table 4. All factor loadings were equal to or above the recommended cutoff value of 0.32 (Tabachnick and Fidell (2012).

Test–retest reliability

We ran bivariate correlations between Time 1 and Time 2 for the CCQ. Regarding the morningness-eveningness dimension, the test–retest reliability was good for Dutch (r = 0.89, $p < 0.001$) and Portuguese (r = 0.89, $p < 0.001$), acceptable for Spanish (r = 0.76, $p < 0.001$) and Italian (r = 0.76, $p < 0.001$), but questionable for Arabic (r = 0.64, $p < 0.001$) and German (r = 0.62, $p < 0.001$). Regarding the amplitude dimension, the test–retest reliability was good for Dutch (r = 0.81, $p < 0.001$), Portuguese (r = 0.80, $p < 0.001$) and Spanish (r = 0.80, $p < 0.001$), acceptable for Arabic (r = 0.71, $p < 0.001$) and Italian (r = 0.78, $p < 0.001$), but unacceptable for German (r = 0.32, $p < 0.001$).

Convergent validity

There were no MEQ data for Arabic and Portuguese samples. For other languages, the MEQ showed negative correlations with the CCQ-ME, with small to medium effect sizes: Dutch (r = –0.45, $p < 0.001$), German (r = –0.22, $p < 0.001$), Italian (r = –0.14, p = 0.003) and Spanish (r = –0.47, $p < 0.001$). The negative correlations can be explained by high scores on the MEQ reflecting a morningness orientation, whereas high scores on the CCQ-ME reflect an eveningness orientation. For the CCQ-A dimension, the MEQ showed trivial to small effect size correlations: German (r = –0.11, p = 0.053), Italian (r = 0.00, p = 0.986), Dutch (r = –0.13, p = 0.003) and Spanish (r = –0.18, p = 0.001).

Criterion validity

Because measurement invariance could not be assumed across the validated versions, the CCQ-ME and CCQ-A scores could not be pooled across languages. Therefore, we explored associations between CCQ scores and theoretical correlates within samples. Findings from these correlation analyses are provided in the Supplementary File. Overall, the data indicate that CCQ-ME had a small–medium negative correlation with age (r values ranged from –0.10 to –0.34),

but the CCQ-A showed no clear pattern of correlations with age. No clear pattern of results emerged for physical activity, shift-work, sleep duration, medical treatment for sleep or coffee consumption. Four of the six samples showed a small positive correlation between the CCQ-A and napping (indicating that a higher amplitude was associated with having a daytime nap), but the CCQ-ME was unrelated to napping.

Discussion

The aim of this study was to translate the CCQ to six languages (namely, Arabic, Dutch, Italian, German, Portuguese and Spanish) and test the validity and test–retest reliability of the measure within each of these languages. Our findings showed that the two-factor structure of the original French version could be replicated in the six languages. However, measurement invariance could not be assumed regarding the factor loadings. Test–retest reliability varied considerably between the six versions of the CCQ ranging from unacceptable reliability (German version) to good reliability (Dutch, Portuguese and Spanish). Convergent validity was established through correlations between the MEQ and the morningness-eveningness dimension of the CCQ. Some criterion validity was established as participants taking daytime naps tended to show greater diurnal fluctuations (CCQ-A), and younger participants tended to show a greater preference for eveningness (CCQ-ME).

Regarding the factor structure of the six language adaptations, all samples showed a good fit to the theoretically predicted two-factor structure, confirming the findings of the original version in the French language (Dosseville et al. 2013). However, measurement invariance could not be assumed across versions due to discrepancies related to factor loadings. The difference in factor loadings could have several causes. First, there were some notable differences in demographic characteristics between samples. We endeavored to be as homogeneous as possible in our data collection, but could not avoid sample particularities that might have influenced the measurement invariance test between languages. For example, in the Arabic sample more than half of the sample declared doing shift work, while shift work was almost absent from the German sample. Further, the German sample was quite young in comparison to the other samples, and displayed more than twice as much physical activity hours per week than the other samples. Similarly, taking a nap was reported by more than 40% of the Spanish sample but less than 10% of the German sample. These demographic differences might explain the different factor loadings across the different versions of the CCQ. In addition to differences in sample characteristics, differences stemming from the environment might also affect chronotype scores, such as sunrise time, sunset time and light exposure (see Figueiro et al. 2014; Leocadio-Miguel et al. 2017). That some items load higher or lower on the factors according to the language means that the factor scores obtained cannot be directly compared across languages, and that their interpretation requires some care. A similar measurement invariance analysis should be realized before any cross-cultural comparison related to chronotype is made in order to ensure the validity of the comparison. This unfortunately has not been the case in previous cross-cultural studies involving the CCQ and other chronotype questionnaires (e.g. Randler et al. 2015).

Test–retest reliability was investigated to assess the reliability of the CCQ. Test–retest reliability was not assessed in the original validation of the CCQ (Dosseville et al. 2013) and therefore offers added value to the current study. Regarding the CCQ-ME dimension, test–retest reliability was good for Dutch and Portuguese, acceptable for Italian and Spanish, and questionable for German and Arabic. Regarding the CCQ-A dimension, test–retest reliability was good for Dutch, Portuguese and Spanish, acceptable for Arabic and Italian, and unacceptable for German. The unacceptable test–retest reliability of the CCQ-A in the German sample might be explained by some particularities of the sample. The German sample were largely undergraduate (sport science) students who are required to take part in a range of sport classes in which times cannot be chosen. This might lead to amplitude variations across time that might differ between Time 1 and Time 2. The effects of training and change in fitness levels between Time 1 and Time 2 might also have led to some natural changes in scores on the amplitude dimension of the CCQ. Nevertheless, more research is required on the

German version of the questionnaire to check on the reliability of the measure across time.

This study also tested convergent validity of the CCQ-ME with the MEQ (Horne and Ostberg 1976). Convergent validity was established, given that a negative association emerged between the MEQ and the CCQ-ME for the four languages tested (the direction of the coefficient explained by the MEQ morningness being at the higher end of the continuum). The correlation values observed ranged from small to medium. This is somewhat lower than in the original validation study that observed a large effect size correlation between the two questionnaires, $r = -0.82$ (Dosseville et al. 2013), and might reflect aforementioned sample particularities. For the CCQ-A dimension, either a trivial association (for the German and Italian samples) or small association (for the Dutch and Spanish samples) with the MEQ was observed. This reflects to some extent the findings of the original validation study, with $r = -0.21$ (Dosseville et al. 2013). This negative correlation might indicate that the MEQ mixes partly amplitude information with morningness-eveningness information, but could also reflect the small–medium positive correlation between the amplitude and morningness-eveningness dimensions that is often observed in research (see Supplementary File). In other words, people with a tendency toward eveningness also show (somewhat) larger diurnal fluctuations (Baehr et al. 2000).

There was also some evidence of criterion validity as dimensions of the CCQ appeared to relate to daytime napping and participant age. It was not possible to combine the samples for a single analysis due to measurement invariance, but notable trends across samples were evident. First, a small-medium negative correlation emerged in all samples between participant age and the CCQ-ME, indicating that younger participants showed a greater tendency towards eveningness (older participants had a greater tendency toward morningness). This finding is in line with previous research showing that chronotype tends toward morningness after a peak in eveningness during late adolescence or early adulthood (Fischer et al. 2017; Randler et al. 2017; Roenneberg et al. 2004). Second, four of the six samples showed that napping had a small positive correlation with the CCQ-A, indicating that people who take daytime naps show greater diurnal fluctuations (higher amplitude). This finding complements previous research that has found some evidence for an association between chronotype and napping (e.g. Lee et al. 2017; Suh et al. 2017), and makes intuitive sense given that napping lowers physiological activation (and therefore contributes to greater variability). Nevertheless, as this was not observed in two of the six samples, further research is required to verify this observation.

The current study has some potential limitations that require careful consideration from readers when interpreting main findings. First, the samples included were opportunistic and are not necessarily representative of the population (nation) as a whole. Second, the data were collected over a long period of time and differences between samples might be due to seasonal variations that we were unable to control for. Third, the English translation of the CCQ (on which all translations are based) has still not undergone a comprehensive validation procedure. Fourth, several sleep variables related to morningness-eveningness, such as mid-sleep, habitual and preferred bedtimes, and rise times on weekdays and weekends (Oginska, 2011; Randler et al. 2016) have not been considered in this study, and further research should take them into account to further establish the validity of the CCQ-ME. Fifth, after the CCQ data collection started, a new questionnaire was created to assess morningness-eveningness and amplitude: the Morningness-Eveningness-Stability-Scale-improved (MESSi, Randler et al. 2016; Diaz-Morales et al. 2017), and this new questionnaire should be compared to the CCQ in future research. Sixth, the test–retest reliability period of 1 month was not only selected based on previous research, but also needs to be examined over a longer time period to further establish scale reliability (Kantermann & Eastmann 2018). Finally, while the MEQ (Horne and Ostberg 1976) has been validated and compared to physiological measures, this is still missing for the CCQ. Examples of physiological data that could be assessed alongside the CCQ are body temperature variation during the day (Baehr et al. 2000) and melatonin onset (Burgess and Fogg 2008). Validation against physiological markers will be particularly useful to test the quartile cutoff values presented in Table 2, and identify whether high and low scorers on the CCQ-ME and CCQ-A can also be distinguished based on physiological data. We hope that the translation of the CCQ into several languages will provide

opportunities for researchers to assess the physiological correlates of chronotype across world regions.

To conclude, this study provides evidence of construct validity for the CCQ across six languages. However, measurement invariance could not be assumed regarding the factor loadings. Convergent validity was established through correlations between the CCQ and an established measure of morningness-eveningness, and criterion validity was provided in the form of positive correlations between age and morningness-eveningness, and between daytime napping and amplitude. There was mixed evidence for the test–retest reliability of the CCQ across languages, and further work testing the reliability of the questionnaire is recommended. Overall, our findings generally support the use of the translated versions of the CCQ to investigate chronotype across world regions. We recommend researchers explore how dimensions of the CCQ relate to physiological markers of health and well-being. Specifically, assessing chronotype amplitude seems particularly relevant for domains such as sport, work and health organizations, and research about chronotype amplitude might help to prevent a broad range of psychological and physiological disorders, chronic fatigue and other sleep issues (see Dosseville et al. 2013).

Acknowledgements

Sylvain Laborde and Fabrice Dosseville coordinated the validation project of the Caen Chronotype Questionnaire and wrote the first draft of the paper. All coauthors provided their feedback on the paper, and then Mark Allen finalized the manuscript with Sylvain Laborde. The coauthors are listed by alphabetical order. We indicate below the researchers responsible for each version:

Arabic version: Karim Chamari coordinated the translation and the data collection of the Arabic version, with the help of Asma Aloui, Helmi Ben Saad, Hamdi Chtourou, Abdulaziz Farooq, Monoem Haddad and Karim Khalladi. We would like to thank as well May Awartani and Fawaz Hamie for their help with the translation.

Dutch version: Marijke C.M. Gordijn, Yvonne A.W. de Kort and Karin C.H.J. Smolders coordinated the translation and the data collection of the Dutch version.

German version: Sylvain Laborde coordinated the translation and the data collection of the German version; Barbara Braun; Thomas Hosang and Alexander Wurm helped with the final data analysis.

Italian version: Maurizio Bertollo, Laura Bortoli and Claudio Robazza coordinated the translation and the data collection of the Italian version.

Portuguese version: Pablo Greco and Mariana Lopes coordinated the translation and the data collection of the Portuguese version.

Spanish version: Félix Guillén coordinated the translation and the data collection of the Spanish version.

Declaration of interests

The author(s) declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID

Sylvain Laborde http://orcid.org/0000-0003-4842-6548
Helmi BEN SAAD http://orcid.org/0000-0002-7477-2965
Karim Chamari http://orcid.org/0000-0001-9178-7678
Monoem Haddad http://orcid.org/0000-0001-5989-1627
Claudio Robazza http://orcid.org/0000-0002-3639-1539

References

Adan A, Archer SN, Hidalgo MP, Di Milia L, Natale V, Randler C. 2012. Circadian typology: A comprehensive review. Chronobiol Int. 29(9):1153–75. doi:10.3109/07420528.2012.719971. PMID: 23004349.

Aschoff J, Pohl H. 1978. Phase relations between a circadian rhythm and its zeitgeber within the range of entrainment. Naturwissenschaften. 65(2):80–84. doi:10.1007/BF0044054. PMID: 345129.

Baehr EK, Revelle W, Eastman CI. 2000. Individual differences in the phase and amplitude of the human circadian temperature rhythm: With an emphasis on morningness-eveningness. J Sleep Res. 9(2):117–27. doi:10.1046/j.1365-2869.2000.00196.x. PMID: 10849238.

Burgess HJ, Fogg LF. 2008. Individual differences in the amount and timing of salivary melatonin secretion. PLoS One. 3(8):e3055. doi:10.1371/journal.pone.0003055. PMID: 18725972.

Byrne B. 2009. Structural equation modeling with AMOS: Basic concepts, applications, and programming. New York (NY): Routledge.

Cheung GW, Rensvold RB. 2002. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling. 9(2):233–55.

Cohen D. 1988. Statistical power analysis for behavioral sciences. New York, NY: Routledge Academic.

Coogan AN, McGowan NM. 2017. A systematic review of circadian function, chronotype and chronotherapy in attention deficit hyperactivity disorder. Atten Defic

Hyperact Disord. 9(3):129–47. doi:10.1007/s12402-016-0214-5. PMID: 28064405.

Diaz-Morales JF, Randler C, Arrona-Palacios A, Adan A. 2017. Validation of the MESSi among adult workers and young students: General health and personality correlates. Chronobiol Int. 34(9):1288–99. doi:10.1080/07420528.2017.1361437. PMID: 28933566.

Dosseville F, Laborde S, Lericollais R. 2013. Validation of a Chronotype questionnaire including an amplitude dimension. Chronobiol Int. 30(5):639–48. doi:10.3109/07420528.2012.763042. PMID: 23510464.

Ethnologue. 2018. Languages of the world. Dallas (Texas): SIL International [accessed 2018 Mar 16] https://www.ethnologue.com

Figueiro MG, Plitnick B, Rea MS. 2014. The effects of chronotype, sleep schedule and light/dark pattern exposures on circadian phase. Sleep Med. 15(12):1554–64.

Fischer D, Lombardi DA, Marucci-Wellman H, Roenneberg T. 2017. Chronotypes in the US - Influence of age and sex. PLoS One. 12(6):e0178782. doi:10.1371/journal.pone.0178782. PMID: 28636610.

Hambleton RK, De Jong JHL. 2003. Advances in translating and adapting educational and psychological tests. Lang Test. 20(2):127–34.

Horne JA, Ostberg O. 1976. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol. 4(2):97–110. doi:10.3109/07420529709001458. PMID: 1027738.

Hu L, Bentler P. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling. 6(1):1–55.

Jöreskog KG. 1993. Testing structural equation models. In Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park (CA): Sage Publications; p. 294–316.

Kantermann T, Eastman CI. 2018. Circadian phase, circadian period and chronotype are reproducible over months. Chronobiol Int. 35(2):280–88.

Laborde S, Guillen F, Dosseville F, Allen MS. 2015. Chronotype, sport participation, and positive personality-trait-like individual differences. Chronobiol Int. 32(7):942–51. doi:10.3109/07420528.2015.1055755. PMID: 26181469.

Lee TY, Chang PC, Tseng IJ, Chung MH. 2017. Nocturnal sleep mediates the relationship between morningness-eveningness preference and the sleep architecture of afternoon naps in university students. PLoS One. 12(10):e0185616. doi:10.1371/journal.pone.0185616. PMID: 29040275.

Leocadio-Miguel MA, Louzada FM, Duarte LL, Areas RP, Alam M, Freire MV, Fontenele-Aroujo J, Menna-Barreto L, Pedrazzoli M. 2017. Latitudinal cline of chronotype. Sci Rep. 7(1):5437.

Malone SK, Patterson F, Lozano A, Hanlon A. 2017. Differences in morning-evening type and sleep duration between Black and White adults: Results from a propensity-matched UK Biobank sample. Chronobiol Int. 34(6):740–52. doi:10.1080/07420528.2017.1317639. PMID: 28488939.

Merikanto I, Lahti T, Puolijoki H, Vanhala M, Peltonen M, Laatikainen T, Vartiainen E, Salomaa V, Kronholm E, Partonen T. 2013. Associations of chronotype and sleep with cardiovascular diseases and type 2 diabetes. Chronobiol Int. 30(4):470–77. doi:10.3109/07420528.2012.741171. PMID: 23281716.

Mota MC, Waterhouse J, De-Souza DA, Rossato LT, Silva CM, Araujo MB, Tufik S, De Mello MT, Crispim CA. 2016. Association between chronotype, food intake and physical activity in medical residents. Chronobiol Int. 33(6):730–39.

Nowakowska-Domagala K, Mokros L, Jablkowska-Gorecka K, Grzelinska J, Pietras T. 2016. The relationship between chronotype and personality among patients with alcohol dependence syndrome: pilot study. Chronobiol Int. 33(10):1351–58. doi:10.1080/07420528.2016.1213738. PMID: 27715334.

Nunally JC. 1978. Psychometric theory. New York (NY): McGraw-Hill.

Ogińska H. 2011. Can you feel the rhythm? A short questionnaire to describe two dimensions of chronotype. Pers Individ Dif. 50(7):1039–43.

Ogińska H, Mojsa-Kaja J, Mairesse O. 2017. Chronotype description: In search of a solid subjective amplitude scale. Chronobiol Int. 34(10):1388–400.

Partonen T. 2015. Chronotype and health outcomes. Current Sleep Medicine Reports. 1(4):205–11.

Portaluppi F, Smolensky MH, Touitou Y. 2010. Ethics and methods for biological rhythm research on animals and human beings. Chronobiol Int. 27(9–10):1911–29. doi:10.3109/07420528.2010.516381. PMID: 20969531.

Randler C, Diaz-Morales JF, Rahafar A, Vollmer C. 2016. Morningness-eveningness and amplitude - development and validation of an improved composite scale to measure circadian preference and stability (MESSi). Chronobiol Int. 33(7):832–48. doi:10.3109/07420528.2016.1171233. PMID: 27123963.

Randler C, Fassl C, Kalb N. 2017. From Lark to Owl: Developmental changes in morningness-eveningness from new-borns to early adulthood. Sci Rep. 7:45874. doi:10.1038/srep45874. PMID: 28378787.

Randler C, Prokop P, Sahu S, Haldar P. 2015. Cross-cultural comparison of seven morningness and sleep-wake measures from Germany, India and Slovakia. Int J Psychol. 50(4):279–87. doi:10.1002/ijop.12098. PMID: 25175755.

Randler C, Stadler L, Vollmer C, Diaz-Morales JF. 2012. Relationship between depressive symptoms and sleep duration/chronotype in women. J Individ Differ. 33(3):186–91.

Refinetti R, Lissen GC, Halberg F. 2007. Procedures for numerical analysis of circadian rhythms. Biol Rhythm Res. 38(4):275–325.

Roenneberg T, Kuehnle T, Pramstaller PP, Ricken J, Havel M, Guth A, Merrow M. 2004. A marker for the end of adolescence. Curr Biol. 14(24):R1038–1039. doi:10.1016/j.cub.2004.11.039

Seo DG, De Jong G. 2015. Comparability of online- and paper-based tests in a statewide assessment program. J Educ Comput Res. 52(1):88–113.

Simor P, Zavecz Z, Palosi V, Torok C, Koteles F. 2015. The influence of sleep complaints on the association between chronotype and negative emotionality in young adults. Chronobiol Int. 32(1):1–10. doi:10.3109/07420528.2014.935786. PMID: 25003651.

Suh S, Yang HC, Kim N, Yu JH, Choi S, Yun CH, Shin C. 2017. Chronotype differences in health behaviors and health-related quality of life: A population-based study among aged and older adults. Behav Sleep Med. 15 (5):361–76. doi:10.1080/15402002.2016.1141768. PMID: 27148632.

Tabachnick B, Fidell L. 2012. Using multivariate statistics. 6th ed. New Jersey, USA: Pearson.

Tsaousis I. 2010. Circadian preferences and personality traits: A meta-analysis. Eur J Pers. 24(4):356–73.

Von Schantz M, Taporoski TP, Horimoto AR, Duarte NE, Vallada H, Krieger JE, Pedrazzoli M, Negrao AB, Pereira AC. 2015. Distribution and heritability of diurnal preference (chronotype) in a rural Brazilian family-based cohort, the Baependi study. Sci Rep. 5:9214. doi:10.1038/srep09214. PMID: 25782397.

Whittier A, Sanchez S, Castaneda B, Sanchez E, Gelaye B, Yanez D, Williams MA. 2014. Eveningness chronotype, daytime sleepiness, caffeine consumption, and use of other stimulants among peruvian university students. J Caffeine Res. 4(1):21–27.

Wickwire EM, Geiger-Brown J, Scharf SM, Drake CL. 2017. Shift work and shift work sleep disorder: Clinical and organizational perspectives. Chest. 151(5):1156–72. doi:10.1016/j.chest.2016.12.007. PMID: 28012806.