










# Developing the Premonitory Urges for Tic Disorders Scale—Revised (PUTS-R)

Leoni Baumung<sup>1</sup>, Kirsten Müller-Vahl<sup>2</sup> , Katherine Dyke<sup>3\*</sup> ,  
Georgina Jackson<sup>4,5</sup> , Stephen Jackson<sup>3,5</sup> , Dennis Golm<sup>6</sup> ,  
Alexander Münchau<sup>1</sup>  and Valerie Brandt<sup>6</sup> 

<sup>1</sup>Department of Paediatric and Adult Movement Disorders and Neuropsychiatry, Institute of Neurogenetics, Center for Brain, Behavior and Metabolism, University of Lübeck, Germany

<sup>2</sup>Clinic of Psychiatry, Socialpsychiatry and Psychotherapy, Hannover Medical School, Germany

<sup>3</sup>School of Psychology, University of Nottingham, UK

<sup>4</sup>Division of Psychiatry and Applied Psychology, School of Medicine, University of Nottingham, UK

<sup>5</sup>Institute of Mental Health, School of Medicine, University of Nottingham, UK

<sup>6</sup>Department of Psychology, Centre for Innovation in Mental Health, University of Southampton, UK

**Background.** Patients with Gilles de la Tourette syndrome (GTS) or chronic tic disorders frequently experience premonitory urges prior to tics. The 'Premonitory Urges for Tic Disorders Scale' (PUTS) is commonly used in order to assess urge severity in patients with tics. Several studies suggest that the PUTS might measure more than one dimension of urges. These include the quality and severity of premonitory urges.

**Methods.** This study aims to replicate and extend previous findings concerning the psychometric properties of the PUTS and its underlying dimensions in a large sample of 241 patients with GTS including both adults ( $n = 93$ ; mean age =  $34.2 \pm 12.84$ ; 73 male) and minors ( $n = 148$ ; mean age =  $11.8 \pm 2.86$ ; 123 male), pooled from three different recruitment sites.

**Results.** Data analysis confirmed good reliability across the PUTS items for both minors and adults and acceptable item characteristics for items 2–8. A factor analysis of items 1–8 confirmed the existence of two factors in both age groups.

**Conclusions.** The results suggest that the PUTS might benefit from several further small modifications, such as rephrasing items 1 and 9 to increase convergence with the overall construct of the scale. Finally, we propose a revised version of the PUTS, consisting of two subscales: one for *urge severity* and another one for *urge quality* by including several new items.

---

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

\*Correspondence should be addressed to Katherine Dyke, School of Psychology, University of Nottingham, Nottingham NG7 2RD, UK (email: [katherine.dyke@nottingham.ac.uk](mailto:katherine.dyke@nottingham.ac.uk)).

The vast majority of adolescents and adults with Gilles de la Tourette syndrome (GTS) experience 'premonitory urges' or simply 'urges,' immediately prior to a tic. The prevalence of these urges has been reported to occur in approximately 77% of GTS patients over 13 years old, and in 90% of those over 18 years (Bliss, 1980; Brandt, Beck, Sajin, Baaske, *et al.*, 2016; Kwak, Dat Vuong, & Jankovic, 2003; Sambrani, Jakubovski, & Muller-Vahl, 2016). Awareness of these feelings of premonitory urge tends to increase in children as they become older (Banaschewski, Woerner, & Rothenberger, 2003), potentially as a consequence of the normal development of self-awareness/bodily urges. For most, these unpleasant, involuntary sensations are commonly accompanied by a feeling of unease or anxiety which can be relieved by executing a tic (Brandt, Beck, Sajin, Baaske, *et al.*, 2016; Cavanna & Nani, 2013; Kwak *et al.*, 2003). Premonitory urges are a core feature of GTS for many, and awareness of them is critical for behavioural interventions such as habit reversal therapy (HRT) (Azrin & Nunn, 1973) and extensions of this, such as comprehensive behavioural intervention for tics (CBIT) (Piacentini *et al.*, 2010). Hence, to enhance our understanding of GTS and improve treatment options, there is a clear need to develop reliable and accurate approaches to measuring premonitory urges.

The most commonly used approach to assessing urges in GTS is through the use of the Premonitory Urges for Tics Scale (PUTS). This scale was developed by Woods, Piacentini, Himle, and Chang (2005) and is a short, easily administered questionnaire. PUTS has been shown to have good to acceptable reliability in individuals over 10 years old and good convergent validity (Brandt, Beck, Sajin, Anders, & Munchau, 2016). However, internal consistency in children younger than 11 years appears to be questionable (Raines *et al.*, 2018; Steinberg *et al.*, 2010; Woods *et al.*, 2005). Correlations in the low–medium range between tic severity scores and the PUTS are in line with the assumption that these two phenomena are related but distinct constructs (Brandt, Beck, Sajin, Anders, *et al.*, 2016; Ganos *et al.*, 2012; Raines *et al.*, 2018). However, substantial positive correlations of the PUTS with obsessive–compulsive symptoms and anxiety (Rajagopal & Cavanna, 2014; Reese *et al.*, 2014; Steinberg *et al.*, 2010; Woods *et al.*, 2005) point towards low discriminant validity of at least some of the PUTS items.

Originally, the PUTS questionnaire was designed to assess a one-dimensional construct, that is premonitory urges. However, factor analyses suggest that the PUTS measures at least two dimensions (Brandt, Beck, Sajin, Anders, *et al.*, 2016; Raines *et al.*, 2018): (1) the intensity or frequency of urge phenomena and (2) the sensory quality of urges. Hence, the development of different PUTS subscales might be useful. Furthermore, studies investigating changes in premonitory urge intensity after treatment have often failed to show any differences when this was assessed using the PUTS (e.g., Houghton *et al.*, 2017; Nonaka *et al.*, 2015). While it is possible that urges are not affected by treatment, the findings could also be due to a lack of sensitivity in the measure. Specifically, relatively few items of the scale may measure urge intensity and these items may be too vague to assess subtle changes.

The main aim of the current study is to confirm and extend previous findings regarding item characteristics, reliability and underlying dimensions of the PUTS both in minors and adults with GTS, in a sample large enough to produce robust results. To our knowledge, this study also represents the first to investigate psychometric characteristics of the PUTS, based on the item response theory (IRT). The second aim of the study is to use the results to make clear recommendations for a revised version of the PUTS and to develop an item pool that can be tested to develop an urge scale based on more favourable psychometric properties. In doing so, this research is the first step in developing a revised PUTS.

## Materials and methods

### **Participants and clinical assessments**

The data used in this study represent secondary analysis of PUTS scores, which were collected as part of routine assessment along with the Yale Global Tic Severity Scale (YGTSS) during a range of different experiments performed across three sites.

The study included 241 patients with a confirmed diagnosis of a tic disorder according to DSM-IV-TR (DSM-IV-TR, 2000) or DSM-5 criteria (DSM-5, 2013), of which 93 were adults (mean age =  $34.2 \pm 12.84$ ; 73 male) and 148 were minors under the age of 18 years (mean age =  $11.8 \pm 2.86$ ; 123 male). At the time of assessment,  $n = 72$  adults fulfilled criteria for GTS,  $n = 1$  for a chronic phonic tic disorder and  $n = 20$  for a chronic motor tic disorder.  $N = 131$  minors fulfilled criteria for GTS,  $n = 3$  for a chronic phonic tic disorder and  $n = 14$  for a chronic motor tic disorder. Of the minors, 98 were between the ages of 11 and 18 (mean age =  $13.5 \pm 1.9$ ; 76 male) and 50 were 10 years or younger (range = 6–10; mean age =  $8.6 \pm 1.23$ , 50 male). Data were pooled from three different sites and across several experiments: Hannover, Germany ( $n = 96$ ; collected 2013–2015) and Lübeck, Germany ( $n = 80$ ; collected 2011–2017) and Nottingham, UK ( $n = 65$ ; 2013–2016). Of all the studies, these data were pooled from, only one with a sample size of 15 excluded participants based on comorbidity. All other studies included a representative sample of GTS participants – some of whom had comorbidities or were taking medication. All patients and parents, respectively, had given their written informed assent/consent prior to taking part in the primary study. The data were anonymized before pooling. Each primary study was reviewed and approved by the respective local ethics committee and conformed to the Declaration of Helsinki.

### **Questionnaires**

All patients filled out the PUTS (Woods *et al.*, 2005), a 10-item questionnaire, assessing urge intensity on a 4-point Likert scale (range = 10–40). Small children received help from their parents filling out the PUTS. Parents were instructed as follows: ‘please help your child fill out this questionnaire. Please explain to your child what each question means and discuss with them which response might be most accurate. If you have any questions, please ask the experimenter. German centres used the translated and validated German version of the PUTS’ (Rössner, Müller-Vahl, & Neuner, 2010). In addition to the complete PUTS scale (referred to as PUTS<sub>1–10</sub>), we explore the 9-item PUTS score (PUTS<sub>1–9</sub>) because item 10 (‘I am able to stop my tics even if only for a short period of time’) is commonly dropped from the overall questionnaire score (Reese *et al.*, 2014; Woods *et al.*, 2005).

The Yale Global Tic Severity Scale (YGTSS) was also completed by 236/241. The YGTSS is a structured and validated interview which measures total tic severity (TTS; range = 0–50) during the past 7 days, with good reliability (Storch *et al.*, 2005). Total tic severity is calculated from the sum of the subscores which measure the number, frequency, intensity, complexity and interference of motor and phonic tics. A separate item which is not factored into TTS assesses overall impairment and is scored out of 50. Subscores were available for 69/93 adults and 142/148 minors for further analyses.

### **Statistical analysis**

Of all PUTS values, 1% were missing. Missing values were not replaced because the rate was <5% (Bennett, 2001; Schafer, 1999). Item difficulty and discrimination were tested

according to the classical test theory (CTT) in SPSS (IBM Corp, 2016) and the IRT in R (R Development Core Team, 2017).

Item difficulty indicates how many participants have answered an item above the mean of the scale. For items with multiple response options, item difficulty is calculated as follows:  $p = \sum_i \frac{x_i}{k \cdot n}$  (Bortz & Döring, 2013; Moses, 2017). Values for each item are summed up across participants ( $x_i$ ) and are divided by the number of participants ( $n$ ) multiplied by the item response levels ( $k$ ). Items that all participants answer with '4' or '1' are not useful because they do not discriminate between participants. In CCT, item difficulty can range between 0 and 1; items with a difficulty under 0.2 or over 0.8 are commonly excluded from questionnaires. In IRT, a similar parameter is called 'threshold' for items with multiple response options. Thresholds ( $b$  values) indicate the point of ability or underlying construct where participants switch from one response option to the next. If the item measures the intensity of an underlying construct, the probability of switching to a higher response option should be ordered according to the response option (i.e., patients with very high urge intensities should be more likely to select 4 as a response option than 3). Non-ordered thresholds indicate that patients with higher symptom severity do not necessarily select higher values on this particular item and that the item might therefore not be ideal to measure the construct of the scale.

Item discrimination refers to the ability of an item to discriminate between participants who will score high or low on a questionnaire, that is how well one item reflects the whole scale. In CTT, this is assessed with part-whole corrected item-to-total correlation (Pearson's  $r$ , values  $>.40$  are considered adequate). In IRT, items with slope values of 0.65–1.34 are considered moderate, values  $>1.34$  high (R Development Core Team, 2017).

Cronbach's alpha was calculated as a measure of reliability (internal consistency) for minors and adults, respectively, and for minors  $\leq 10$  and  $>10$  respectively. Cronbach's alpha  $>.80$  is considered good. Correlations between the PUTS items and the YGTSS scores were conducted using Spearman's rank coefficients.

Furthermore, a weighted least squares factor analysis for ordinal data with varimax rotation was conducted in Lisrel (Jöreskog & Sörbom, 2018). Due to low inter-item correlations (Table S1), items 9 and 10 were not included in the factor analysis (Field, 2013; Raines *et al.*, 2018).

### **Item development**

To extend the *urge quality* subscale, items were developed based on the literature (Banaschewski *et al.*, 2003; Kurlan, Lichter, & Hewitt, 1989). Every quality described in the literature in association with urges was included as an item, using the phrasing of the original PUTS '*right before I do a tic...*'. Existing items that contained more than one construct (item 1 refers to a ticklish or itchy feeling) were rephrased into separate items, so that each item refers to one construct only (Price, Jhangiani, & Chiang, 2015).

For the severity items, a deductive approach was used (Burisch, 1984). The chosen definition of the construct to be measured was developed based on the pre-existing literature and reads as follows: 'premonitory urges are uncomfortable sensations or feelings that increase before a tic is executed but may vary in intensity between individuals, tics and time points.' Based on this construct, VB and AM formulated a set of items, based on the principle that items should be 'brief,' 'relevant,' 'unambiguous,' 'specific' and 'objective' (Price *et al.*, 2015). The items were reviewed and adjusted after feedback from KMV, KD, GJ and SJ. The questionnaire was then translated to German by VB and back-translated to English by DG. English and German versions of the

questionnaire were given to two patients each for initial feedback regarding the non-ambiguity and face validity of the items.

It is important to stress that these items will now need evaluation and selection based on their psychometric properties. The items should be given to a large group of TS patients, together with a number of measures that assess the same construct (i.e., urge severity) and related but different constructs (e.g., obsessive-compulsive disorder [OCD]) in order to determine individual item properties, convergent validity (high correlations with measures that assess the same construct) and discriminant validity (low-medium correlations with scales that assess other, possibly related constructs). Factor analyses should be used to test dimensionality of questionnaires. Items with unfavourable psychometric properties should be excluded from the final version of a revised PUTS scale.

## Results

### **Clinical assessment**

In minors (<18 years), the average PUTS<sub>1-10</sub> score was  $21.08 \pm 6.32$  and YGTSS-TTS was  $20.20 \pm 9.18$ . In adults, the average PUTS<sub>1-10</sub> score was  $23.84 \pm 6.25$  and the YGTSS-TTS was  $20.34 \pm 9.18$ . Independent samples t tests revealed no statistically significant differences in YGTSS-TTS scores between minors and adults,  $t(234) = -0.12, p = .904$ . Adults had significantly higher PUTS<sub>1-10</sub> scores than minors,  $t(239) = 3.31, p = .001$ .

### **Item difficulty and discrimination**

Item 10 shows high item difficulty for both adults and minors, indicating that most participants select 3 or 4 for this item. Hence, this item is not ideally suited to differentiate between patients with intense and less intense urges. Non-ordered thresholds further indicate that patients with less intense urges do not necessarily select less intense response options for items 9 and 10. The same is true for item 5 in adults and item 1 in both groups. Item 1 has a rather low item difficulty, indicating that most patients select 1 or 2 and only few patients select higher response options. Item 10 in both groups and item 1 in adults had low item-test correlations, indicating they did not reflect the construct measured by the rest of the scale (see Table 1).

### **Reliability**

Cronbach's alpha across the PUTS<sub>1-10</sub> items was good in minors ( $n = 136$ )  $\alpha = .82$  and acceptable in adults ( $n = 88$ )  $\alpha = .78$ . Cronbach's alpha was better across the PUTS<sub>1-9</sub> in minors ( $n = 137$ )  $\alpha = .83$  as well as adults ( $n = 89$ ),  $\alpha = .82$ . Reliability of the PUTS<sub>1-9</sub> was also good in children younger than 11 years ( $n = 45$ ;  $\alpha = .83$ ).

### **Dimensions of the PUTS**

Varimax rotated factor solutions across PUTS<sub>1-8</sub> items for minors and adults, respectively, showed two factors (see Table 2). Items 1, 6, 7 and 8 loaded on one factor (previously termed intensity items), items 2-5 (adults) and 3-5 (minors) loaded on one factor (quality items). Items 2 loaded on different underlying factors in minors and the adults.

**Table 1.** Item difficulty, thresholds and discrimination of the 10 PUTS items

Item-test correlations minors										
	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10
<b>CTT</b>										
PUTS <sub>1-10</sub>	.40***	.58***	.53***	.56***	.53***	.47***	.68***	.54***	.40***	.19*
Item difficulty	.38	.48	.50	.46	.41	.55	.63	.52	.69	.72
$\alpha$ if item removed	.81	.79	.80	.79	.79	.80	.78	.79	.81	.83
<b>IRT</b>										
Factor corr	0.41	0.6	0.45	0.46	0.46	0.42	0.85	0.68	0.31	0.1
a	0.77	1.27	0.86	0.88	0.87	0.78	2.73	1.57	0.56	0.17
b1	1.34	0.02	-0.07	0.31	1.1	-0.15	-0.81	0.14	-0.29	-3.79
b2	2.22	1.03	1.27	1.32	1.38	0.52	0.11	0.31	-1.27	0.6
b3	1.30	1.31	1.08	1.34	1.41	1.10	0.73	1.25	0.20	-3.63
<b>Adults CTT</b>										
PUTS <sub>1-10</sub>	.33***	.49***	.52***	.65***	.50***	.52***	.60***	.49***	.44***	.00
Item difficulty	.40	.63	.64	.53	.45	.70	.63	.52	.70	.79
$\alpha$ if item removed	.78	.76	.76	.74	.76	.76	.74	.76	.77	.82
<b>IRT</b>										
Factor corr	0.24	0.42	0.5	0.67	0.46	0.54	59	0.46	0.29	<0.01
a	0.42	0.78	1.00	1.55	0.87	1.09	1.24	0.89	0.52	-0.01
b1	1.89	-0.38	-0.38	0.08	0.83	-1.22	-0.62	-0.13	-0.99	214.9
b2	4.16	-0.04	-0.3	0.26	0.72	-0.16	-0.34	0.84	-0.92	116.2
b3	-0.12	0.43	0.46	1.02	1.88	0.07	1.02	1.06	0.23	62.17

Note. Item difficulty and discrimination (CTT: corrected item-to-total correlation; IRT: a) of the 10 PUTS items. Part-whole corrected item-test correlations show that all items, apart from item 10, have a significant moderate association with the rest of the PUTS across both groups. Item difficulty is shown to be moderate for most items and across both groups. However, items 9 and 10 had a very high item difficulty. Factor correlations indicate the correlation of each item with the underlying factor of the scale; values should >.40. Thresholds (b1, b2, b3) indicate the point where participants switch from one response option to the next, that is patients with more intense urges should choose higher values. Values in the acceptable range are black, other values are grey.

CTT = classical test theory; IRT = item response theory.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

**Table 2.** Varimax rotation across PUTS<sub>1-10</sub> items

Items: Right before I do a tic. . .	Minors		Adults	
	Intensity	Quality	Intensity	Quality
<i>Item 1</i> –...I feel like my insides are itchy	<b>0.53</b>	0.17	<b>0.32</b>	0.17
<i>Item 2</i> – . . .I feel pressure inside my brain or body	<b>0.58</b>	0.45	<b>0.27</b>	<b>0.59</b>
<i>Item 3</i> – . . .I feel ‘wound up’ or tense inside	0.32	<b>0.67</b>	0.27	<b>0.75</b>
<i>Item 4</i> – . . .I feel like something is not ‘just right’	0.23	<b>0.80</b>	0.29	<b>0.87</b>
<i>Item 5</i> –...I feel like something is not complete	0.22	<b>0.80</b>	0.19	<b>0.65</b>
<i>Item 6</i> –...I feel like there is energy in my body that needs to get out	<b>0.48</b>	0.36	<b>0.53</b>	0.46
<i>Item 7</i> – I have these feelings almost all the time before I do a tic	<b>0.94</b>	0.27	<b>0.92</b>	0.23
<i>Item 8</i> – These feeling happen for every tic I have	<b>0.78</b>	0.20	<b>0.71</b>	0.22

Note. The results of the varimax rotation across PUTS<sub>1-8</sub> items in both groups. The factors ‘intensity’ and ‘quality’ are shown for each group. Values shown in bold indicate on which factor each variable loaded the highest.

### Convergent validity between the PUTS and YGTSS scale

Spearman’s rank correlations between the PUTS factors identified above and YGTSS subscales in minors and adults can be found in Table 3.

## Discussion

### Psychometric properties of the PUTS

The primary aim of this study was to examine the psychometric properties according to CTT and IRT, and dimensions of the PUTS in minors and adults with GTS in a large sample. The second aim was to use these results to make clear recommendations for a revised PUTS. Key findings are summarized below.

**Table 3.** Correlations PUTS factors and YGTSS subscales

	Minors		Adults	
	Intensity	Quality	Intensity	Quality
Motor				
Number	<b>.30 (&lt;.001)</b>	<b>.18 (.036)</b>	.03 (.779)	<b>.32 (.008)</b>
Frequency	<b>.29 (.001)</b>	<b>.22 (.009)</b>	.15 (.224)	–.004 (.977)
Intensity	<b>.24 (.004)</b>	<b>.23 (.007)</b>	.18 (.114)	.15 (.225)
Complexity	<b>.17 (.038)</b>	.08 (.355)	.20 (.104)	<b>.39 (.001)</b>
Interference	<b>.18 (.035)</b>	<b>.28 (&lt;.001)</b>	.20 (.095)	<b>.24 (.043)</b>
Phonic				
Number	<b>.17 (.043)</b>	.14 (.107)	.05 (.704)	–.03 (.826)
Frequency	<b>.21 (.012)</b>	<b>.21 (.011)</b>	.08 (.522)	.07 (.576)
Intensity	.15 (.078)	<b>.18 (.03)</b>	.13 (.295)	.06 (.619)
Complexity	<b>.21 (.013)</b>	<b>.24 (.004)</b>	.20 (.105)	.14 (.262)
Interference	.08 (.335)	.13 (.13)	.16 (.177)	.14 (.243)
Impairment	<b>.31 (.003)</b>	<b>.30 (.004)</b>	.12 (.312)	.22 (.073)

Note. Spearman’s  $\rho$  correlations between the previously identified PUTS factors and the YGTSS subscales for minors and adults respectively. Values shown in bold indicate statistically significant results.

### **Item difficulty and discrimination**

Combining CTT and IRT, the statistical analyses confirm that item 10 of the PUTS (*suppressibility of tics*) has unacceptable thresholds/ item difficulty and discrimination parameters in minors and adults. The findings are in line with previous studies (Brandt, Beck, Sajin, Anders, *et al.*, 2016; Reese *et al.*, 2014; Woods *et al.*, 2005).

Item response theory-based analyses showed that thresholds were not ordered for items 1 and 9. Previous studies have already shown small inter-item (Raines *et al.*, 2018) and item-test correlations (Brandt, Beck, Sajin, Anders, *et al.*, 2016) between items 1 and 9 and the other items of the PUTS as well as another urge measure (Brandt, Beck, Sajin, Anders, *et al.*, 2016). Overall, the results consistently suggest that the psychometric properties of these items are not ideal to assess urges (Brandt, Beck, Sajin, Anders, *et al.*, 2016; Raines *et al.*, 2018). Interestingly, thresholds indicated that item 5 should also be reviewed in the adult PUTS.

The results presented here support previous work which strongly suggests that the PUTS scale would benefit from the removal of item 10 (Capriotti, Brandt, Turkel, Lee, & Woods, 2014; Steinberg *et al.*, 2010). Furthermore, items 1 and 9 should be rephrased in order to enhance their representation of the overall construct measured by the PUTS.

### **Dimensions of the PUTS**

Previous research findings pointed towards three underlying dimensions (Brandt, Beck, Sajin, Anders, *et al.*, 2016) of the PUTS<sub>1-10</sub> or two dimensions, if items 9 and 10 were excluded (Raines *et al.*, 2018). The current study confirms the existence of two distinct dimensions within the PUTS<sub>1-8</sub> in both adults and minors.

In keeping with previous findings, one factor included items 2–5 in adults (Item 2: *feeling pressure*, item 3: *feeling wound up or tense*, item 4: *'not just right' feelings*, item 5: *feeling incompleteness*) and items 3–5 in minors and can be called 'quality of premonitory sensations.'

A second cluster consisted of items 6 (*feeling of energy*), 7 (*'I have these feelings almost all the time before I do a tic'*) and 8 (*feelings happen for every tic*). This factor was originally termed the 'intensity factor' because it loaded on one factor with an independent measure (the 'real-time urge monitor') that assessed urge intensity over time (Brandt, Beck, Sajin, Anders, *et al.*, 2016). However, it has since been suggested that the factor may be more accurately referred to as an 'urge frequency' factor (Raines *et al.*, 2018). It is possible that urge intensity as measured by the real-time urge monitor (Brandt, Beck, Sajin, Baaske, *et al.*, 2016) is highly correlated with urge frequency. Correlations with the YGTSS subscales did indeed show that the intensity/ frequency PUTS factor had its highest correlations with the YGTSS tic frequency and tic number subscales in minors, closely followed by tic intensity, while it seemed to have its highest correlations with tic complexity in adults. It is therefore unclear whether this factor represents the same underlying construct in minors and adults. Furthermore, it is unclear how well urge intensity and frequency can be distinguished. An additional limiting factor is that the results do not only depend on the psychometric properties of the PUTS but also the YGTSS'. As far as the authors are aware, the subscales of the YGTSS have also never been confirmed using factor analysis. Further research will be needed with instruments that can differentiate between urge intensity and frequency. We will therefore refer to the intensity/ frequency PUTS factor as *urge severity* for the moment.



### **Reliability and validity**

One of the main criticisms levelled at the PUTS is a that it has previously been found to have poor psychometric properties in children 10 years and younger (Martino *et al.*, 2017). However, in contrast to previous studies (Raines *et al.*, 2018; Steinberg *et al.*, 2010; Woods *et al.*, 2005), our work revealed good internal consistency in children younger than 11. It is currently not clear why this is the case, but it is possible that having help from their parents increased reliability of the questionnaire. Future studies should determine the factors that increase reliability of the PUTS in young children in order to optimize study settings (e.g., filling out questionnaires together with a parent or clinician, at home or in the clinic).

### **The PUTS revised**

Based on these findings, we would like to propose a number of changes to the PUTS scale. In general, we would like to propose two PUTS subscales: (1) an *urge severity* subscale and (2) an *urge quality* subscale (Appendix). Given the relatively small pool of patients with GTS in a given place, we would like to encourage colleagues in the field to help us evaluate and select items for a revised PUTS either by running their own validation studies or by getting in contact with us to collaborate, using our protocol.

The developers of the PUTS have already suggested to drop item 10 from the overall PUTS score (Reese *et al.*, 2014; Woods *et al.*, 2005); in the same vein, we would suggest to drop the item altogether if adjustments are made to revise the scale.

Regarding the structure of the scale, we would propose two changes: (1) change the scale to a 5-point Likert scale so that the scale has a mid-point (please see Appendix for a suggestion of the revised PUTS) and (2) the scale should range from 0 to 4 so that no symptoms or not agreeing with the item corresponds to the value 0 instead of 1. Furthermore, not all items should be phrased so that a higher number represents more severe urges. This helps participants to pay attention to the scale and can help identify participants who automatically tick the same response for each item, without reading the items.

Regarding the content of the PUTS, we would suggest testing several additional items using statistical techniques to establish a questionnaire with two dimensions (*urge severity* and *urge quality*) in which only items with good or excellent psychometric properties are included. Based on the existing literature, our expertise and the results of this study, we have suggested some additional items and rephrased some existing ones, in order to increase the reliability of the *urge severity* subscale (Appendix). Future research should investigate the reliability, validity and dimensions of the current version of the revised PUTS scale and exclude unsuitable items in order to create a final PUTS-R scale with excellent psychometric properties.

Regarding the *urge quality* subscale, we would propose to add more items, based on the literature regarding the different qualitative descriptions of urges (Banaschewski *et al.*, 2003; Kurlan *et al.*, 1989).

Future studies might compare and contrast comorbidities, such as OCD and attention-deficit hyperactivity disorder (ADHD) and their possible relationship with different *urge quality* items in patients with tics (Brandt, Beck, Sajin, Anders, *et al.*, 2016). Whether the *urge quality* items should be included in a total PUTS score or whether the two subscales should be viewed as entirely separate scales should also be investigated, using other measures of urge intensity and frequency in order to test the validity of the scale. Alternatively, the severity subscale might be a useful clinical indicator of urge intensity or

frequency, while the quality subscale of premonitory urges might be viewed as qualitative information for clinicians and a research tool, rather than a severity index. Previous research suggests that there might be interesting associations between premonitory urges in general and obsessive–compulsive behaviour – specifically ‘just-right’ experiences (Sambrani *et al.*, 2016). This association that might be captured by a particular PUTS quality item asking about just-right experiences (Brandt, Beck, Sajin, Anders, *et al.*, 2016) and should be further investigated.

It should be noted that another version of the PUTS, the individualized PUTS (I-PUTS), exists (McGuire *et al.*, 2016). This questionnaire assesses presence, frequency, intensity and body region of urges for each tic that a patient reports (symptom checklist parallel to the YGTSS). A strength of this questionnaire is that it is likely sensitive to change. Weaknesses include that it was not developed based on psychometric properties, and some psychometric properties have been assessed but convergent validity was low and that its length and score depends on the number of tics a patient reports.

### **Limitations**

Comorbidities and the influence of medication were not analysed in the current study. The factor structure was not tested in children <10 years independently because this group was not large enough. Whether the PUTS is suitable for children aged 10 and younger needs to be further explored.

### **Conflicts of interest**

All authors declare no conflict of interest.

### **Author contributions**

Leoni Baumung (Formal analysis; Writing – original draft); Kirsten Müller-Vahl (Conceptualization; Data curation; Investigation; Project administration; Writing – review and editing); Katherine Dyke, Ph.D (Data curation; Investigation; Project administration; Validation; Writing – review and editing); Georgina Jackson (Conceptualization; Data curation; Investigation; Project administration; Validation; Writing – review and editing); Stephen Jackson (Conceptualization; Data curation; Investigation; Project administration; Validation; Writing – review and editing); Dennis Golm (Investigation; Validation; Writing – review and editing); Alexander Münchau (Conceptualization; Data curation; Investigation; Project administration; Writing – review and editing); Valerie Brandt (Conceptualization; Data curation; Methodology; Project administration; Supervision; Validation; Writing – review and editing).

### **Data availability statement**

The data that support the findings of this study are available on request from the corresponding author, if data sharing agreement is signed. The data are not publicly available due to ethical restrictions.

## References

- Azzin, N. H., & Nunn, R. G. (1973). Habit-reversal: A method of eliminating nervous habits and tics. *Behavioural Research and Therapy*, *11*, 619–628. [https://doi.org/10.1016/0005-7967\(73\)90119-8](https://doi.org/10.1016/0005-7967(73)90119-8)
- Banaschewski, T., Woerner, W., & Rothenberger, A. (2003). Premonitory sensory phenomena and suppressibility of tics in Tourette syndrome: Developmental aspects in children and adolescents. *Developmental Medicine and Child Neurology*, *45*, 700–703. <https://doi.org/10.1017/S0012162203001294>
- Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, *25*, 464–469. <https://doi.org/10.1111/j.1467-842X.2001.tb00294.x>
- Bliss, J. (1980). Sensory experiences of Gilles de la Tourette syndrome. *Archives of General Psychiatry*, *37*, 1343–1347. <https://doi.org/10.1001/archpsyc.1980.01780250029002>
- Bortz, J., & Döring, N. (2013). *Forschungsmethoden und evaluation [Research methods and evaluation]* (2nd ed.). Berlin, Heidelberg, Germany: Springer Verlag.
- Brandt, V. C., Beck, C., Sajin, V., Anders, S., & Munchau, A. (2016). Convergent validity of the PUTS. *Frontiers in Psychiatry*, *7*, 51. <https://doi.org/10.3389/fpsy.2016.00051>
- Brandt, V. C., Beck, C., Sajin, V., Baaske, M. K., Baumer, T., Beste, C., . . . Munchau, A. (2016). Temporal relationship between premonitory urges and tics in Gilles de la Tourette syndrome. *Cortex*, *77*, 24–37. <https://doi.org/10.1016/j.cortex.2016.01.008>
- Burisch, M. (1984). Approaches to personality inventory construction. *A Comparison of Merits. American Psychologist*, *39*, 214–227. <https://doi.org/10.1037/0003-066X.39.3.214>
- Capriotti, M. R., Brandt, B. C., Turkel, J. E., Lee, H. J., & Woods, D. W. (2014). Negative reinforcement and premonitory urges in youth with Tourette syndrome: An experimental evaluation. *Behavior Modification*, *38*, 276–296. <https://doi.org/10.1177/0145445514531015>
- Cavanna, A. E., & Nani, A. (2013). Tourette syndrome and consciousness of action. *Tremor and Other Hyperkinetic Movements*, *3*, <https://doi.org/10.7916/D8PV6J33>
- DSM-5. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- DSM-IV-TR. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (4th ed.). Washington, DC: American Psychiatric Association.
- Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). London, UK: Sage.
- Ganos, C., Kahl, U., Schunke, O., Kuhn, S., Haggard, P., Gerloff, C., . . . Munchau, A. (2012). Are premonitory urges a prerequisite of tic inhibition in Gilles de la Tourette syndrome? *Journal of Neurology, Neurosurgery and Psychiatry*, *83*, 975–978. <https://doi.org/10.1136/jnnp-2012-303033>
- Houghton, D. C., Capriotti, M. R., Scahill, L. D., Wilhelm, S., Peterson, A. L., Walkup, J. T., . . . Woods, D. W. (2017). Investigating habituation to Premonitory Urges in Behavior Therapy for Tic Disorders. *Behavior Therapy*, *48*, 834–846. <https://doi.org/10.1016/j.beth.2017.08.004>
- IBM Corp. (2016). *IBM SPSS statistics for windows, version 24.0*. Armonk, NY: Author.
- Jöreskog, K. G., & Sörbom, D. (2018). *LISREL 9.30 for windows [Computer software]*. Skokie, IL: Scientific Software International Inc.
- Kurlan, R., Lichter, D., & Hewitt, D. (1989). Sensory tics in Tourette's syndrome. *Neurology*, *39*, 731–734. <https://doi.org/10.1212/WNL.39.5.731>
- Kwak, C., Dat Vuong, K., & Jankovic, J. (2003). Premonitory sensory phenomenon in Tourette's syndrome. *Movement Disorders*, *18*, 1530–1533. <https://doi.org/10.1002/mds.10618>
- Martino, D., Pringsheim, T. M., Cavanna, A. E., Colosimo, C., Hartmann, A., Leckman, J. F., . . . Martinez-Martin, P. (2017). Systematic review of severity scales and screening instruments for tics: Critique and recommendations. *Movement Disorders*, *32*, 467–473. <https://doi.org/10.1002/mds.26891>
- McGuire, J. F., McBride, N., Piacentini, J., Johnco, C., Lewin, A. B., Murphy, T. K., & Storch, E. A. (2016). The premonitory urge revisited: An individualized premonitory urge for tics scale.

- Journal of Psychiatric Research*, 83, 176–183. <https://doi.org/10.1016/j.jpsychires.2016.09.007>
- Moses, T. (2017). A review of developments and applications in item analysis. In R. E. Bennett & M. von Davier (Eds.), *Advancing human assessment. Methodology of educational measurement and assessment* (pp. 19–46). Cham, Switzerland: Springer.
- Nonaka, M., Matsuda, N., Kono, T., Fujio, M., Scahill, L., & Kano, Y. (2015). Preliminary study of behavioral therapy for Tourette syndrome patients in Japan. *Children's Health Care*, 44, 293–306. <https://doi.org/10.1080/02739615.2014.979922>
- Piacentini, J., Woods, D. W., Scahill, L., Wilhelm, S., Peterson, A. L., Chang, S., . . . Walkup, J. T. (2010). Behavior therapy for children with Tourette disorder: A randomized controlled trial. *Journal of the American Medical Association*, 303, 1929–1937. <https://doi.org/10.1001/jama.2010.607>
- Price, P. C., Jhangiani, R., & Chiang, I. (2015). *Research Methods of Psychology* (2nd ed.). Victoria: BCcampus. Retrieved from <https://opentextbc.ca/researchmetho>.
- R Development Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Raines, J. M., Edwards, K. R., Sherman, M. F., Higginson, C. I., Winnick, J. B., Navin, K., . . . Specht, M. W. (2018). Premonitory Urge for Tics Scale (PUTS): Replication and extension of psychometric properties in youth with chronic tic disorders (CTDs). *Journal of Neural Transmission (Vienna)*, 125, 727–734. <https://doi.org/10.1007/s00702-017-1818-4>
- Rajagopal, S., & Cavanna, A. E. (2014). Premonitory urges and repetitive behaviours in adult patients with Tourette syndrome. *Neurological Sciences*, 35, 969–971. <https://doi.org/10.1007/s10072-014-1706-8>
- Reese, H. E., Scahill, L., Peterson, A. L., Crowe, K., Woods, D. W., Piacentini, J., . . . Wilhelm, S. (2014). The premonitory urge to tic: Measurement, characteristics, and correlates in older adolescents and adults. *Behavior Therapy*, 45, 177–186. <https://doi.org/10.1016/j.beth.2013.09.002>
- Rössner, V., Müller-Vahl, K., & Neuner, I. (2010). PUTS – premonitory urge tics scale: Fragebogen für Kinder [PUTS – premonitory urge tics scale: questionnaire for children]. In K. Müller-Vahl (Ed.), *Tourette-Syndrom und andere Tic Erkrankungen im Kindes und Erwachsenenalter [Tourette syndrome and other tic disorders in childhood and adolescence]* (p. 204). Berlin, Germany: MWV Medizinische Wissenschaftliche Verlagsgesellschaft.
- Sambrani, T., Jakubovski, E., & Muller-Vahl, K. R. (2016). New insights into clinical characteristics of Gilles de la Tourette Syndrome: Findings in 1032 patients from a Single German Center. *Frontiers in Neuroscience*, 10, 415. <https://doi.org/10.3389/fnins.2016.00415>
- Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, 8(1), 3–15. <https://doi.org/10.1191/096228099671525676>
- Steinberg, T., Shmuel Baruch, S., Harush, A., Dar, R., Woods, D., Piacentini, J., & Apter, A. (2010). Tic disorders and the premonitory urge. *Journal of Neural Transmission (Vienna)*, 117, 277–284. <https://doi.org/10.1007/s00702-009-0353-3>
- Storch, E. A., Murphy, T. K., Geffken, G. R., Sajid, M., Allen, P., Roberti, J. W., & Goodman, W. K. (2005). Reliability and validity of the Yale Global Tic Severity Scale. *Psychological Assessment*, 17, 486–491. <https://doi.org/10.1037/1040-3590.17.4.486>
- Woods, D. W., Piacentini, J., Himle, M. B., & Chang, S. (2005). Premonitory Urge for Tics Scale (PUTS): Initial psychometric results and examination of the premonitory urge phenomenon in youths with Tic disorders. *Journal of Developmental and Behavioral Pediatrics*, 26, 397–403. <https://doi.org/10.1097/00004703-200512000-00001>

## Supporting Information

The following supporting information may be found in the online edition of the article:

**Table S1.** Item intercorrelations for minors and adults.

## Appendix:

### *Premonitory Urge for Tics Scale – Revised (PUTS-R)*

Below, you will find a list of statements concerning feelings that can occur together with tics. Please indicate how much you agree with each statement from 0 (not at all) to 4 (very much). The statements refer to your average/ typical feelings. When you respond to the statements, please think about typical everyday situations. Please respond to every statement. If you do not agree with a statement or you do not experience the described feeling in association with tics, please tick 0.

Name/ ID \_\_\_\_\_ Age \_\_\_\_\_ Place: school clinic home other

Date \_\_\_\_\_ Diagnosis (if known) \_\_\_\_\_

How I feel	Not at all				Very much				
	0	1	2	3	4				
1	Right before I do a tic, I have a general unpleasant or strange feeling that I cannot describe in more detail								
2	Right before I do a tic I feel pressure inside my brain or body								
3	Right before I do a tic I feel 'wound up' or tense inside								
4	Right before I do a tic I feel like something is not 'just right'								
5	Right before I do a tic I feel like something isn't complete								
6	Right before I do a tic I feel like there is energy in my body								
7	Right before I do a tic I have a feeling that is similar to an itch								
8	Right before I do a tic I have a feeling that is similar to a ticklish sensation								
9	Right before I do a tic, I have a feeling that is similar to a cold sensation								
10	Right before I do a tic, I have a feeling that is similar to a warm sensation								
11	Right before I do a tic, I have a feeling that is similar to a burning sensation								
12	Right before I do a tic, I have a feeling that is similar to numbness inside my head or body								
13	Right before I do a tic, I have a feeling like a shiver or twitch in my head or body								
14	Right before I do a tic, I have a feeling that can be compared to something touching my body								
15	Right before I do a tic, I have a feeling that is similar to an ache in my head or body								
16	I have these feelings almost all the time before I do a tic								

*Continued*

**Appendix.** (Continued)

---

Name/ ID \_\_\_\_\_ Age \_\_\_\_\_ Place: school clinic home other  
 Date \_\_\_\_\_ Diagnosis (if known) \_\_\_\_\_

---

How I feel	Not at all				Very much
	0	1	2	3	4
17 I rarely have these feelings before a tic					
18 These feelings happen for every tic I do					
19 These feelings are strong					
20 This feeling gets stronger until I do a tic					
21 Most of my tics just happen, I don't feel them coming					
22 When I have these feelings, I cannot think of anything else but to do a tic					
23 I have these feelings only before some tics					
24 Overall, how strong were these feelings in the past week?					

---

*Notes.* The current version of the PUTS-R contains items of the PUTS and additional suggestions for items developed by experts in the field (urge severity) and based on a literature search (urge quality). These items should be assessed based on a large sample. Only items with good or excellent psychometric properties should be used for the final PUTS-R scale, other items should be discarded.