Anspannung

Müller, Friedrich; Mori, Shuji; Sakaki, Yuko; Yi, Kwangoh; Bae, Sungbong; Tan, Yuka; Ward, Lawrence M.

Published in:
Methods in Psychology

DOI:
10.1016/j.metip.2022.100098

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Methods in Psychology 7 (2022) 100098

Contents lists available at ScienceDirect

Methods in Psychology

journal homepage: www.sciencedirect.com/journal/methods-in-psychology

Anspannung: Introduction to concept and quantification of mental strain exemplified on data taken in five countries

Friedrich Müller [a,*], Shuji Mori [b], Yuko Sakaki [c], Kwangoh Yi [d], Sungbong Bae [d], Yuka Tan [e], Lawrence M. Ward [f]

[a] Institute of Experimental Industrial Psychology, Leuphana University of Lüneburg, Wilschenbrucher Weg 84a, D-21335, Lüneburg, Germany
[b] Department of Informatics, Faculty of Information Science and Electrical Engineering, Kyushu University, Japan
[c] Department of Human Sciences, Chikushi Jogakuen University, Japan
[d] Department of Psychology, Yeungnam University, South Korea
[e] Department of Applied Japanese, Southern Taiwan University of Science and Technology, Taiwan
[f] Department of Psychology and Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Canada

ARTICLE INFO

Keywords:
Anspannung
Category partitioning
Cultural comparisons
Mental strain
Psychological tension
Psychophysical scaling

ABSTRACT

In Dükers action theory *Anspannung*, which we translated as "psychological tension" (PT), is described as a directly experienced valid indicator for the extent of mental strain. In German-speaking regions the Category Partitioning technique (CP) has proven to be a useful method for accurately quantifying the experienced PT. Outside Germany, however, the concept of PT and the CP technique for measuring it have found little resonance, as it seemed that the central terms could not be meaningfully translated into English. To challenge these language barriers, test the applicability and usefulness of the PT concept, and evaluate the CP scaling method, we used the CP technique to quantify the level of PT required by 32 imagined everyday situations. To do this we adapted descriptions of the everyday situations from the German into English, Japanese, Korean, and Mandarin Chinese, and enrolled $N = 158$ participants from five countries (Canada, Germany, Japan, South Korea, and Taiwan). The results show a remarkable agreement between the data collected in the five cohorts. The experimental data point to the universality of the experience of mental load in culturally and linguistically diverse societies. They also point to the need to design scaling techniques so that respondents can describe their immediate sensations as they would in everyday life.

1. Objective of this study

The category partitioning technique (CP) introduced here is a scaling method developed to measure the intensity of experienced attributes. Since it aims at the quantitative description of the phenomenal world, the method reacts sensitively to the formulation of the questioned properties. This was particularly evident when scaling mental strain, whose experiential component is described as *Anspannung* in German, by searching for equivalent terms in other languages. The exploratory studies described here were intended to examine whether measures of the experienced psychological tension that would accompany imagined everyday activities are comparable among people with different cultural and linguistic backgrounds.

2. Anspannung or psychological tension

2.1. The phenomenon

In his theory of action H. Düker (1975) defines four directly experienceable factors, which are decisive for the realization of actions. These are the interest (in), the striving (for), the action-adequate mental activation and the decision to act (see also Tent, 1988). From this follows that every state of a person and every activity performed is linked with a specific degree of activation. Bartenwerfer (1969) termed this unspecific but indispensable excitation for the initiation and successful execution of actions "General Central Activation" (GCA, p. 195) (German: *Allgemeine zentrale Aktiviertheit*). The more demanding an activity is, the higher the level of activation must be. GCA is an indicator of mental strain, defined in EN ISO 10075 as "… the immediate effect of

* Corresponding author. Wilschenbrucher Weg 84a, D-21335, Lüneburg, Germany.
E-mail address: f.mueller@leuphana.de (F. Müller).

https://doi.org/10.1016/j.metip.2022.100098
Received 9 September 2021; Received in revised form 12 May 2022; Accepted 16 June 2022
Available online 22 June 2022

mental stress within the individual (not the long-term effect) depending on his/her individual habitual and actual preconditions, including individual coping styles." (p. 1).

The physiological component of GCA can be tapped indirectly by measuring heart rate, heart rate variability, blood pressure, and by analyzing EEG patterns and other biological processes involved in activation. Because these measures are very unspecific, however, it is hardly possible, especially in practical application, to estimate the respective proportions of such measures caused by specific activities.

A more direct access to the activation processes, however, seems achievable by quantifying the *perceived* components of GCA. This is possible because the degree of activation is directly experienced in every moment of waking life. In German-speaking countries, activation is experienced and communicated as *Anspannung* which could be translated as "psychological tension" (PT). It seems we can provide precise information on the degree of PT at any given time. Reports of everyday events often contain information about the level of PT. For example, if a student after completing a psychophysics exam reports: "The test wasn't difficult - I was much less tense than I expected." Or when the racing driver Sebastian Vettel reports after a race in heavy rain: "Das Rennen erforderte maximale Anspannung während des gesamten Verlaufs (This race required maximum PT during the entire course)." If this experience can be measured precisely, PT would be a useful indicator of mental strain. The unique advantage is the unmediated access to the specific phenomenon.

2.2. Quantification of PT

2.2.1. Bartenwerfer scale

Based on scaling results of 557 workers and employees, Bartenwerfer (1963, 1969) developed a numerical sensation scale that covers the number range from 0 to 50. The scale is anchored by written situations: "0 = *deep dreamless sleep*" and "48 = *I am sitting full of the fear of death in a crashing airplane.*" As internal anchors he placed the descriptions of nine further situations at previously determined corresponding scale points (see Fig. 1).

Outside the laboratory, 18 workers of a metallurgical plant and 10 supervising workplace experts scaled with nearly perfect agreement (average difference of scale scores M = 0.9; SD = 0.8) the tension associated with eight different work activities.

And, while simultaneously observing the change in heart rate compared to a base rate occurring during a simple activity (silent counting of knocks), in seven further experiments the participants carried out 17 mental activities of varying complexity. The correlation of heart rate with the PT values assigned to the activities was r = 0.78. From these findings Bartenwerfer (1969) deduced that his PT scale was an interval scale.

Müller (1981) proposed several ways in which Bartenwerfer's scaling approach could be improved. These referred to the anchoring of the scale, the formulation of the anchor situations, and the instructions given to participants. The external anchors do give an idea of the extent of possible degrees of PT. However, since presumably none of the participants have been experienced being in a crashing airplane, both external anchors lie outside of their conscious experience. And, as the level of experienced strain is influenced by practice and familiarity with the situations described the given hierarchy of the internal anchors may confuse the participants in which case the simple description of the immediately experienced PT could be confounded with comparisons of the experience of ego-distant predefined standards.

Therefore, in order to prevent multiple comparisons and evaluations, and to enable an unbiased description of the referenced experience, we required the participants to orient their judgments with respect to self-experienced situations going from the lowest to highest experienced PT.

The same applies to the self-focused formulation (*I* solve crossword puzzles, *I* don't indicate when *I* experience severe pain). This calls for introspective making of comparisons not only to other personal

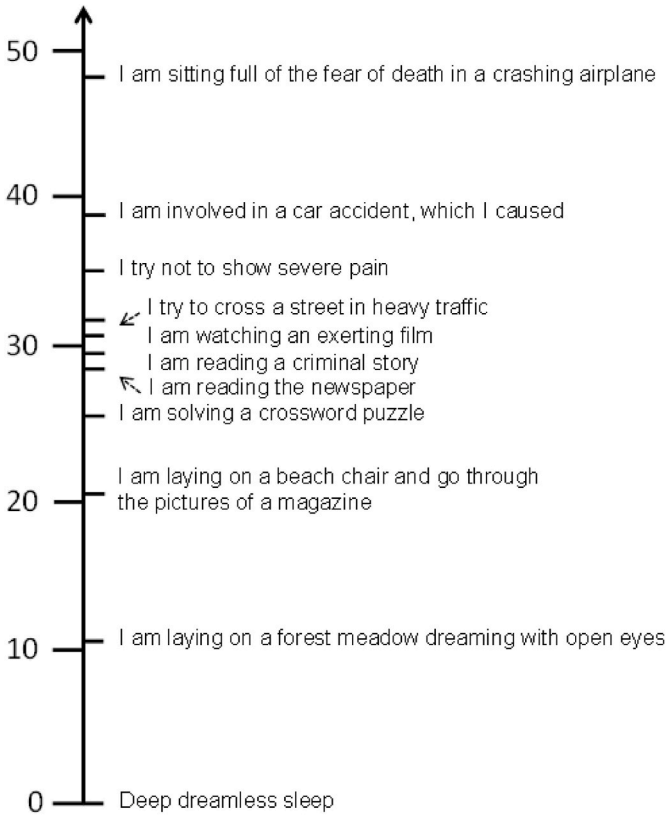

Fig. 1. Sensation scale for measuring psychological tension as used by Bartenwerfer (1969).

experiences but also to more general principles. The appeal to evaluate one's own, very personal experience raises the question, and raises doubts, as to whether one's own evaluation is correct and agrees with the feelings of other people. Uncertainty again invites critical weighing and multiple comparisons with previously scaled situations, which finally affect the metric properties of the resulting scale.

2.2.2. The category partitioning technique

With the aim of resolving this conflict between perception and judgement, Heller (1982) developed the Category Partitioning technique (CP). Theoretical foundations for the development of the method can be found in the reference system theory (*Bezugssystemtheorie*) of the Tübingen School (Witte, 1960a, 1960b) and Heller's orientation concept (Heller, 1980). The main goal of the CP procedure is to ensure an unbiased and unreflective metric description of the perceived world.

Originally developed as a method for differentiated measurement of suprathreshold hearing for individual fitting of digital hearing aids (Heller, 1985, 1991; Heller et al., 1995), the CP method allows precise description of loudness and pitch perception of individuals and even single ears.

The CP method consists of a name-giving scale and a set of measurement rules as a prerequisite for an authentic quantitative description of perceived properties.

Requirements for obtaining true-to-life measurements are:

- Category scales which contain category designations taken from everyday experience and formulated in common everyday language.
- An ecologically valid and isomorphic relation between the range of stimuli and wording and range of the scale used, whereby, if several stimuli are scaled in sequence, the stimulus set should represent the entire range of experience. Establishing an isomorphic relation between scale and stimulus range is a prerequisite to prevent range effects (Müller, 1988; Müller and Steinbach, 1997; Teghtsoonian,

1973). If one follows the reference system theory (Witte, 1960a), the lifelong memorized experiences with differently sized objects of the same class, form and modify so called "mnestically stabilised reference systems" (*mnemisch stabilisierte Bezugssysteme*) (p. 228), which are prerequisites for the unquestioned simultaneous perception of the metric expression going with those objects.

- Allowing fine differentiation within the given scale by a two-step modus (first describing the magnitude by an appropriate verbal category, then finer grading by numerals within the category). This procedure increases the resolution of the scale and reduces so-called frequency effects (Parducci, 1963,).
- Creating an extraspective (in contrast to introspective) scaling mode. Prevent scrutinizing by not asking for "judgments", "guesses" or "impressions". Don't use phrases as "you cannot do wrong". But ask for facts like "how much PT does this task require?" instead of "how tense are you when performing this task?"

Using a sophisticated experimental design, Heller (1981) tested whether scales of PT are metric. For this purpose, he created three lists of about 40 typical situations each from "student life," "driving", and "everyday life at home." Guided by the bisymmetry axiom (Pfanzagl, 1959), over the course of a semester, students were asked several times to pick a situation from one of the lists, whose tension bisects the PTs of each one of the items of the remaining lists. For each of the bisections the items were individually selected for each participant depending on the previously performed bisections. The results confirmed that the CP scaled PT meets the requirements of a metric scale.

As an indication of external validity of the CP scale for PT, Müller (1981) found a linear relation between PT and reaction time in a simple reaction time experiment where variation in PT was induced by the variation of fore-periods. Measured on a verbally anchored 7-point category scale (from very low PT to very high PT), subdivided by 9 graphic markers each, an increase of 1 category goes with a 10.5 ms decrease of reaction time.

2.3. Application in research and practice

PT has also been used in a variety of practical applications. For example, PT as a measure of mental effort and reaction time as a measure of performance were taken in experiments over a 6.5-h period, either in quiet or with continuous exposure to medium loud white noise or an unpleasant complex tone (1000 Hz, 1004 Hz, and 1011 Hz) at the same loudness (Müller, 1992). The results showed that the relation between z-transformed reaction-time and PT, taken as an indicator of the efficiency of mental performance, described a nearly constant effort-outcome relation in quiet and in the white noise condition. The participants working in the presence of the unpleasant noise, described as a "love-crazy cricket," however, became less and less efficient in the course of the experiment.

And in 66-min vigilance tests observing the Mackworth-clock (Mackworth, 1948), as expected by Eysenck and Eysenck (1987), introverted persons, who were identified by the Eysenck personality inventory, reported with a difference of about five CP-units, significantly higher PT than extraverts. However, there were no differences observed in skin conductivity or vigilance (Semjonova, 2012).

PT was also important for the granting of the operating license for high-speed trains (ICE) in Germany. A prerequisite for journeys at speeds of 300 km/h or more is that the strain caused by the trains passing at a distance of about 1.5 m does not cause any risk to the health of the technicians working on the track. Reactions were studied under double blind conditions for speeds of 200 km/h, 280 km/h, and 300 km/h. As indicators for mental strain, PT-scaling as well as recordings of heart rate, inter-beat-intervals, finger pulse amplitude, respiratory rate and variation, EMG of the musculus trapezius, and monitoring the immunoglobulin A concentration in saliva. None of the physiological parameters was sensitive enough to detect differences between conditions. The PT-values, however, showed such a high sensitivity that, in addition to the speed-dependent differences, the course of the pass-by could also be mapped as shown in Fig. 3 (Müller et al., 2003).

2.3. Is anspannung (PT) an exclusively German phenomenon?

Investigations into the experience of PT have so far been limited to the German-speaking world. One reason for this is that the literature on the theoretical and methodological foundations are almost exclusively available in German only. There are only a few English publications that report on the application of the CP method in other sensory domains, such as scaling of pain (Ellermeier et al., 1991), loudness (Hellbrück, 1996),[1] perceived exertion (Kakarot and Müller, 2014; Neely, 1995) or perceived exertion, loudness, and pain (Müller et al., 1995).

Another obstacle could be that there seemed to be no unambiguous translations for some of the concepts important for tension research. One example is the semantic differentiation between *Belastung* (in English: stress, load, strain) and *Beanspruchung* (in English: stress, load, strain), which is central to German stress research (Rohmert, 1984). The English language has a more extensive vocabulary, but the differences in meaning between the terms stress, load, and strain do not seem to be so clearly differentiated in colloquial English.

This linguistic obstacle was ameliorated with the adoption of the German industrial standard DIN ISO 33405 into the British Standard BS EN ISO 10075 (European Committee for Standardization, 2000). There, *Belastung* is now clearly defined as "mental stress" and *Beanspruchung* as "mental strain". For scientific

Communication, stress is defined as "… the total of all assessable influences impinging upon a human being from external sources and affecting it mentally" (p. 1). Mental strain is defined now as "the immediate effect of mental stress on the individual depending on his/her individual habitual and actual preconditions, including individual coping styles" (p. 1).

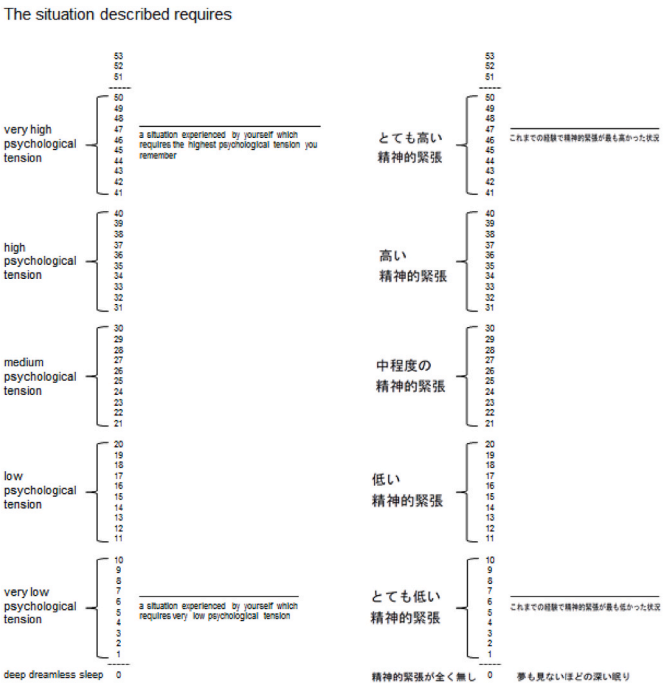

Fig. 2. PT-scales as being used in Canada (left) and in Japan (right).

[1] Instead of CP, Hellbrück used the translation Category-subdivison scaling for the German term Kategorienunterteilungsverfahren

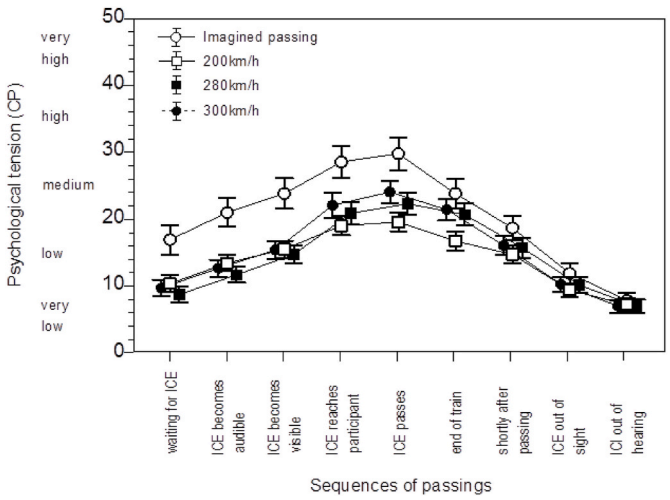

Fig. 3. PT assigned by 48 railroad engineers to different phases of an imaginated train pass-by of an Inter-City Express (ICE) at 300 km/h before the engineers experienced real passings as well as PT measures taken at the same time while real passing with 200 km/h, 280 km/h and 300 km/h (arithmetic means with standard error).

It is even more difficult to find an appropriate translation for the German term *Anspannung* in the meaning of experienced mental strain. When searching for a proper translation, the comment by an English Oxford-educated linguist who was living and teaching in Germany for more than 40 years was: "We don't have an expression for Anspannung and therefore we don't have the phenomenon either," presumably referring to English speakers. This response reflects a strong version of the Sapir-Whorf hypothesis, in which language determines experience. Although many studies have considerably weakened the asserted effects of language on perception and cognition (e.g. Lucy, 1992), the concept of Anspannung seems like an ideal opportunity to contribute further information to this discussion.

As a first step toward the present study, the working group of the research network Perception and Performance in Real and Complex Environments, which was supported by the Japan Society for the Promotion of Science, raised the question of whether it would be possible to find a Japanese equivalent of Anspannung. After several discussion loops and exploratory experiments with Japanese students in Fukuoka and exchange students in Germany, Japanese colleagues proposed the term *sei-sin-te-ki kin-cho*. Experiments conducted in Japan using this description, as reported below, were quite successful. The scaling task was well understood by the student participants and solved without any problems. Since the results of these experiments also corresponded quite well with the data collected in Germany at about the same time, we wondered whether the PT concept could also be applied in other rather similar Asian cultures. We thus conducted tests in South Korea and Taiwan as these societies assume to be highly related in terms of cultural background and language family. In their orthographic systems, Japanese and Chinese share Chinese characters although the Taiwanese use more complicated characters than Japanese. That's why they may understand the other language to a certain extent even if they have not learned the other language. On the other hand, Korea has its own orthographic system, the Hangul letters. But many words originate from Chinese so the Korean and Japanese languages have similarities in pronunciation.

From the viewpoint of linguistic genealogy, it seems that Japanese and Korean belong to the same language family (Murayama, 1963; Tanaka et al., 1975). In linguistic typology, these languages also share common grammatical and morphological characteristics, too (Oe, 1978; Ono, 1980; Umeda, 1982). As Chinese and Japanese are spoken in neighboring regions and share Chinese characters, they have many similarities in imagery and conceptual representations and show a high correlation in word typicality (Yu and Sugimura, 1993).

And again, encouraged by the success in the Asian countries, we finally performed an experiment in English-speaking Canada using the expression "psychological tension" as a surrogate for Anspannung. In what follows we describe these cross-cultural and cross-language experiments investigating the applicability of the concept of Anspannung, or psychological tension.

3. Method

3.1. Participants

The scaling experiments were performed with participants in Canada (CA), Germany (DE), Japan (JP), South Korea (KR) and Taiwan (TW).

Mental strain is influenced by a number of factors, such as age, general and actual condition, level of aspiration, abilities, skills, experience, motivation, attitudes, and others. To minimize these influences, we carried out the scaling experiments with relatively homogeneous groups from the five countries. These were students, 91 women and 67 men, at the participating universities with an average age of 22.4 years (see Table 1).

All participants took part in the experiments out of their own interest. Following the respective rules of the participating universities, we took informed consent from the participants and recorded their name, age, and gender. To avoid German participants suspecting that their personality was being tested in the psychological laboratory, the names of the German participants were neither asked nor otherwise recorded. In addition, their own individual results were given to these participants immediately after the scaling. The overall results only were posted on notice boards.

3.2. Procedure

The experimental procedure is closely adapted from the procedure of Heller (1981). The 32 stimuli (see Table 3) were taken from his list of "everyday life at home", whereby the descriptions of everyday situations, ranging from "shortly before falling asleep" to "through phone being informed by the police of the brother's accidental death", cover the range of possible levels of tension.

The scale is designed in such a way that the two-stage scaling mode is supported, first literally naming the level of tension, then finer gradation in numbers. Fig. 2 shows the scales as they were used in Canada and Japan.

3.2.1. Introducing the PT scale and anchoring

After an orientation about the subject of the study, the participants were told that every state and every activity we perform in waking life goes with a specific level of physical involvement and mental activation, which is reflected in the experience of physical exertion and psychological tension.

If the subjects were interested in participating in the study the scale (Fig. 2) was introduced. When explaining its use, participants were asked to anchor the scale by two definite situations they personally

Table 1
Demographic information on the participants of the present study.

Country	Age		Gender	
	M	*SD*	Women (*n*)	Men (*n*)
All	22.4	2.5	91	67
CA	22.4	2.2	7	3
GER	24.9	2.3	15	17
JP	21.2	0.8	18	12
KR	23.0	2.2	20	20
TW	20.8	2.0	31	15

Table 2
Results of orienting trials.

No	Item			All	DE	JP	KR	TW
				n = 148	n = 32	n = 30	n = 40	n = 46
T1	It is a nice, warm evening in the summer. You are relaxing in a beach chair.	PE	M	4.2	4.6	5.1	**3.6****	**4.0***
			CI	3.7/4.7	4.0/5.2	3.9/6.3	2.9/4.1	2,7/5.3
		PT	M	4.8	4.9	5.2	**4.4***	**4.8**
			CI	4.2/5.4	4.1/5.7	4.1/6.3	3.6/5.3	3.3/6.3
T2	You are working with a percussion drill and are going to drill a hole in the kitchen wall. Accidently you hit the water pipe and water comes out the wall.	PE	M	19.2	20.3	**15.4****	**22.0***	18.5
			CI	17.9/ 20.5	18.4/ 22.2	12.7/ 18.1	19.4/ 24.8	15.8/ 21.2
		PT	M	31.3	32.5	33.8	**29.2***	30.5
			CI	29.9/ 32.4	30.7/ 34.3	31.4/ 36.2	26.5/ 31.4	27.7/ 33.3
T3	Becoming aware of a toddler trying to pull a pot containing boiling water from the stove.	PE	M	12.3	14.0	12.0	13.2	10.5
			CI	11.0/ 13.7	11.1/ 16.9	8.7/15.3	10.4/ 16.3	8.5/12.5
		PT	M	39.0	**35.1****	38.9	39.3	**41.5*****
			CI	37.8/ 40.1	32.1/ 38.1	36.4/ 41.4	37.7/ 40.6	39.6/ 43.4
T4	Carrying a crate containing 12 bottles of water (1 L each) into the 3rd floor using stairs.	PE	M	33.1	**26.7*****	**35.9***	**35.3***	34.0
			CI	31.6/ 34.5	23.5/ 29.9	33.9/ 37.9	33.3/ 37.3	30.8/ 37.2
		PT	M	13.6	**11.9*****	**18.2*****	13.1	12.4
			CI	12.4/ 14.9	9.3/14.5	15.3/ 21.1	10.9/ 15.1	10.0/ 14.8
T5	Lifting a 5 kg-bag from the floor and putting it on a table.	PE	M	21.5	**15.5*****	**25.1***	20.9	23.8
			CI	19.9/ 22.9	13.6/ 17.4	22.0/ 28.2	18.0/ 23.8	20.9/ 26.7
		PT	M	9.5	9.3	**12.4***	8.7	**8.4****
			CI	8.5/ 10.4	8.1/10.5	10.4/ 14.4	7.0/10.4	6.3/10.5
T6	With another person carrying a sofa through a narrow staircase from the 1st to 3rd floor.	PE	M	34.6	35.4	32.8	**35.3****	34.5
			CI	33.3/ 35.7	32.7/ 38.1	30.5/ 35.1	33.0/ 37.1	31.8/ 37.1
		PT	M	21.1	**25.9*****	23.0	**15.9*****	20.9
			CI	19.6/ 22.5	23.5/ 28.3	20.1/ 25.9	13.5/ 18.3	17.9/ 23.9

Note. The physical effort and PT required by the given situations were each quantified using CP scales. The table shows arithmetic means (M) with ranges of 95% confidence intervals (CI). Statistically significant differences, based on sign tests, in each case to the mean of the remaining participants of the other cohorts are read from the binomial distribution, whereby * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0$ 0.001.

experienced. One situation should require extremely low PT and another situation extremely high PT. The freely presented instruction had approximately the following wording:

> "The given scale provides the opportunity to describe the amount of psychological tension required by different tasks or situations. Please imagine two situations which you know from your own experience, one of the situations should require very low PT and the other situation should require the maximum of PT you ever experienced by now. It might be that the situation that you remember will not be accompanied with highest possible PT. If it is so, please put your situations next to the number on the scale which will best represent the intensity".

3.2.2. *Orienting trials*

In order to introduce the participants to the task and at the same time to check whether the subjects are aware of the distinction between muscular effort and mental strain, all but the Canadian participants were introduced to a similar CP-scale for the quantification of physical exertion (PE) and were asked to describe PE as well as PT associated with six training items (T1 to T6) using both scales one after the other.

These six training items are:

T1. On a nice, warm evening in the summer relaxing in a beach chair.
T2. While drilling a hole in the kitchen wall accidently hitting the water pipe. Water comes out of the wall.
T3. Becoming aware of a toddler trying to pull a pot containing boiling water from the stove.
T4. Carrying a crate containing 12 bottles of water (1 L each) into the 3rd floor using stairs.
T5. Lifting a 5 kg-bag from the floor and putting it on a table.
T6. With another person carrying a sofa through a narrow staircase from the 1st to the 3rd floor.

The training situations were chosen as to include items associated with markedly different degrees of physical exertion (where $PE_{T1} < PE_{T5} < PE_{T4}$), associated with markedly different levels of strain (where $PT_{T1} < PT_{T2} < PT_{T3}$), as well as items with markedly different intensities of strain and tension (where $PT_{T4} < PE_{T4}$; $PE_{T2} < PT_{T2}$; $PE_{T3} < PT_{T3}$). The item to be described as the first to provide orientation (T1) should be recognized as having both very low effort and strain. And from the scale values for item T6 it should be recognizable that the additional coordination of the persons involved in the strenuous carrying work requires higher PT than the lifting tasks described by items T4 and T5.

On the basis of the assigned scale values, the investigators could directly check whether the scale values are appropriate to the situation and consistent.

In the case of inconsistent scale values or values strongly deviating from expectations, the phenomena studied and the task was explained to the participants again in different words. Since, however, the scale values of all participants corresponded to our expectations, further explanations were unnecessary in all experiments.

This was the reason why, unlike the procedure in the other countries, the Canadian experimenters used a similar procedure as described above for the introduction to PT scaling in order to familiarize the participants with the scaling of perceived exertion. There, the participants also

Table 3
PT measures for 32 everyday situations collected in five countries.

Item	Pos		All	CA	DE	JP	KR	TW
			$n = 158$	$n = 10$	$n = 32$	$n = 30$	$n = 40$	$n = 46$
Shortly before falling asleep	1	M	3.7	4.0	4.2	4.3	**2.5*****	**4.1***
		CI	3.2/4.3	2.5/5.5	3.5/4.9	2.9/5,7	1.9/3.1	2.6/5.6
Dozing on the balcony while lying in a beach chair on a warm summer evening	29	M	4.7	4.5	**5.7****	5.3	**3.3*****	**5.0****
		CI	4.1/5.3	2.1/6.9	4.8/6.6	4.1/6.5	2.7/3.9	3.5/6.5
Listening to pleasant music while lying in bed	7	M	6.0	6.6	**11.2*****	**4.9****	**4.8*****	**4.1*****
		CI	6.1/7.5	1.9/11.3	9.3/13.1	3.9/5.9	4.0/5.6	2.9/5.3
Teethbrushing	15	M	6.8	5.7	**9.1*****	6.6	8.2	**4.4*****
		CI	6.1/7.5	3.5/7.9	8.0/10.2	5.3/7.9	6.6/9.8	3.4/5.4
Reading the newspaper during breakfast	3	M	10.1	11.9	**14.5*****	8.6	9.9	**7.9*****
		CI	9.1/11.1	6.0/17.8	12.7/ 16.3	7.1/10.1	7.8/12.0	6.2/9.6
Opening the door and taking the newspaper from the paperboy	12	M	11.0	8.1	11.9	12.2	12.3	**9.1***
		CI	10.0/ 12.0	5.2/11.0	10.1/ 13.7	9.9/14.5	10.4/ 14.2	7.3/10.9
Watching the news in television	17	M	13.0	12.4	**16.7***	**11.6***	15.2	**9.7****
		CI	11.8/ 14.3	7.5/17.3	14.4/ 19.0	9.4/13.8	12.5/ 17.9	7.8/11.6
Writing a letter to a close friend	8	M	13.3	15.4	**17.6***	**11.5****	14.2	**10.1*****
		CI	12.1/ 14.4	10.1/ 20.7	14.8/ 20.4	9.3/13.7	12.1/ 16.3	8.5/11.7
Cracking a joke among friends	25	M	13.3	18.0	**21.2*****	**10.7***	**12.5***	**9.2*****
		CI	12.0/ 14.6	11.2/ 24.8	18.6/ 23.8	8.7/12.7	10.9/ 14.1	7.3/11.1
Preparing fried eggs	27	M	13.6	**9.4***	12.4	**13.4***	13.6	15.6
		CI	12.5/ 14.8	6.7/12.1	10.5/ 14.3	10.6/ 16.2	11.7/ 15.5	13.4/ 17.8
Putting on some water for tea	22	M	14.8	9.9	**9.8*****	13.4	**20.1****	15.5
		CI	13.4/ 16.2	3.1/16.2	8.5/11.1	10.6/ 16.2	17.6/ 22.6	12.7/ 18.3
Lining up for lunch in a cafeteria	32	M	15.3	12.5	17.3	14.4	**14.8***	15.6
		CI	14.2/ 16.5	8.4/16.6	15.1/ 19.5	11.6/ 17.2	12.7/ 16.9	13.4/ 17.8
Doing your weekly shopping at a supermarket	16	M	16.0	18.1	**21.0*****	**14.2***	17.7	**11.7*****
		CI	14.7/ 17.2	13.3/ 22.9	18.5/ 23.5	11.9/ 16.5	15.5/ 19.9	9.5/13.9
Playing Monopoly with your family	11	M	16.1	19.7	**19.5****	**11.9*****	18.2	**13.8****
		CI	14.8/ 17.3	15.8/ 23.6	17.4/ 21.6	9.6/14.2	15.9/ 20.5	11.4/ 16.2
Reading a crime thriller while sitting in an armchair	23	M	16.1	19.6	**19.0****	**11.3*****	20.6	**12.7*****
		CI	14.7/ 17.6	12.9/ 26.3	17.1/ 20.9	8.9/13.7	17.5/ 23.7	10.3/ 15.1
Opening a decoratively-wrapped birthday present that was just given to you by a friend	24	M	18.2	19.1	**22.9****	17.2	17.5	15.9
		CI	16.7/ 19.6	13.3/ 24.9	20.6/ 25.2	14.2/ 20.2	14.8/ 20.2	13.0/ 18.8
Trying to end an evangelism conversation with a Jehovah's Witness at the front door while staying friendly	13	M	23.2	22.6	21.4	23.1	**28.2*****	20.4
		CI	22.0/ 24.5	15.4/ 29.8	19.3/ 23.5	20.7/ 25.5	26.4/ 30.0	17.8/ 23.0
Packing your best china for an upcoming move	26	M	23.8	22.5	**17.2*****	26.0	**26.8***	24.8
		CI	22.4/ 25.2	17.3/ 27.7	15.2/ 19.2	23.1/ 20.7	24.5/ 29.1	22.0/ 27.6
Appealing to the plumber over the phone, for the second time, to repair your leaking radiator	30	M	24.2	**31.0***	24.9	23.2	**26.7****	**20.7*****
		CI	22.8/ 25.6	27.3/ 34.7	22.6/ 27.2	20.4/ 26.0	23.9/ 29.5	18.0/ 23.4
Playing chess against an equally-talented opponent	18	M	26.8	**32.9***	27.3	22.3	**28.6***	26.5
		CI	25.4/ 28.2	29.4/ 36.4	24.6/ 30.0	19.1/ 25.5	26.4/ 30.8	23.8/ 29.2
Reading the landlord's letter announcing a 10% rent increase	28	M	28.1	30.3	27.8	25.8	**32.1*****	25.9
		CI	26.9/ 29.4	24.3/ 36.3	25.7/ 29.9	23.3/ 28.3	30.3/ 33.9	23.4/ 28.4
Having to dress and drink coffee hastily because of having overslept	5	M	28.4	28.8	26.5	28.0	27.2	**30.8***
		CI	27.2/ 29.5	24.4/ 33.2	24.1/ 28.9	25.3/ 30.7	25.3/ 29.1	28.5/ 33.1
Upon awaking, realizing that one has overslept	4	M	31.6	30.8	29.2	32.9	31.6	32.7
		CI	30.4/ 32.8	26.9/ 34.7	26.8/ 31.6	30.1/ 35.7	29.7/ 33.5	30.3/ 35.1
Looking for your lottery ticket immediately after the winning numbers are announced thinking you have won $250	31	M	33.2	33.1	33.8	**29.0***	33.3	**35.5****
		CI	31.9/ 34.5	28.5/ 37.7	31.4/ 36.2	25.8/ 32.2	30.9/ 35.7	32.9/ 38.1
Reading the landlord's absolutely unexpected eviction notice	9	M	34.0	34.0	35.1	34.8	**36.7****	30.3
		CI	32.8/ 35.2	29.5/ 38.5	32.9/ 37.3	32.5/ 37.1	35.2/ 38.2	27.5/ 33.1
Discovering that you hit a water main while working with a percussion drill	6	M	34.3	37.8	32.8	36.2	33.8	33.7
		CI						

(*continued on next page*)

Table 3 (*continued*)

Item	Pos		All	CA	DE	JP	KR	TW
			$n = 158$	$n = 10$	$n = 32$	$n = 30$	$n = 40$	$n = 46$
			33.2/	32.0/	30.8/	34.1/	32.1/	31.3/
			35.4	43,6	34.8	38.3	35.5	36.1
Fixing a defective electric main being unsafe whether the wire is live	21	M	37.5	36.8	**35.7***	**38.9***	**40.3*****	35.5
		CI	36.2/	29.7/	34.3/	37.1/	38.0/	32.7/
			38.7	43.9	37.1	40.7	42.6	38.3
Getting stuck between two floors in an elevator at night	14	M	37.7	32.9	34.6	38.4	**38.4***	39.8
		CI	36.5/	28.1/	32.1/	36.0/	36.9/	37.4/
			38.9	37.7	37.1	40.8	39.9	42.2
Breaking a precious Chinese vase while walking through your neighbour's apartment	19	M	37.9	39.4	**33.5*****	39.5	37.5	**39.8*****
		CI	36.9/	35.2/	31.1/	37.8/	36.1/	37.9/
			38.9	43.6	35.9	41.2	38.9	41.7
Being awaked by fire shouts around midnight	10	M	40.6	38.8	**42.3****	40.9	40.2	40.1
		CI	39.6/	32.9/	40.7/	38.9/	38.6/	37.7/
			41.7	44.7	43.9	42.9	41.8	42.5
Becoming aware of a toddler trying to pull a pot containing boiling water from the stove	20	M	40.8	41.6	**37.8****	41.4	**41.4****	**41.7***
		CI	39.7/	37.7/	36.0/	39.4/	39.1/	39.7/
			41.8	45,5	39.6	43.4	43.7	43.7
Being informed by the police through a phone call of your brother's accidental death	2	M	46.8	47.4	46.9	46.9	46.9	46.3
		CI	46.2/	45.1/	45.8/	45.7/	46.2/	45.2/
			47.3	49.7	48.0	48.1	47.6	47.4

Note. Arithmetic mean (M) and ranges of 95% confidence intervals (CI) are listed in ascending order of the overall measures (ALL). In column Pos the position of the situations on the item list is given. Statistically significant differences, based on sign tests in each case to the mean of the remaining participants of the other cohorts, are read from the binomial distribution, whereby * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

selected typical situations that require low and very high exertion. The anchor situations for both scales were then scaled in terms of both PT and exertion. If the scale values adequately depicted the experience of effort/stress, the main trials were started. This modified approach seems to accomplish the goal of orienting participants and checking for understanding equally well. Due to the different approach, however, there is no comparable data for the Canadian participants for the orientation trials.

3.2.3. Main trials

When the orienting trials confirmed that participants had performed appropriately on the PE- and PT-scales, they were asked to assign the corresponding degree of PT to the 32 everyday situations listed in Table 3. For orientation in the item pool, participants were first asked to read the item list, thereby indicating for each item how much tension it requires by noting an abbreviation for the respective verbal category using a pencil.

In a second pass they were asked to write down the appropriate numerical value. The participants were instructed that the quantifications in the second run need not to correspond to the scale in the first run and that they have the possibility, if necessary, to use scale values greater than 50. The freedom in the choice of anchor situations, the PT description by numerical values and the use of the scale seems to be necessary to avoid scrutinizing and comparisons to previously given scale values.

3.3. Statistical Analysis

An initial inspection of the raw data and the descriptive parameters of all data and of the individual country cohorts indicates relatively good agreement in scaling across participants. The mean deviation of scale scores within country groups is SD = 6.52, standard error = 1.3, interquartile range = 9.3, and between cohorts is SD = 7.42, standard error = 0.6, and interquartile range = 10.2.

A comparison of the characteristic values describing the central tendency reinforces the assumptions of Bartenwerfer (1963) and Heller (1981) that PE scales are metric. The arithmetic means (M) calculated for the scaling data for all items in all cohorts are nearly identical to the calculated medians. The mean deviations between M and Median range from 0.02 scale units in the German cohort to 0.35 scale units in the Japanese cohort. The deviations of all medians calculated between the cohorts are symmetrically distributed around the corresponding M with n+ = 97; n- = 94; n0 = 1. Also, rough visual inspection of frequency plots of scale values to the described situations suggest that the data may be normally distributed. However, tests with the Kolmogorov-Smirnov test could not confirm the normal distribution assumption. This is probably due to the unequal allocation of the individual scale values. For example, the scale units in the center and at the border of consecutive categories are used more frequently to describe tension than units about half way between center and border of categories. The most frequently used CP-values in the mid-range are (frequency in brackets): 10(211), 15 (197), 25(201), 30(215), 35(180), 40(164). The least used scale-units were 16(50), 17(58), 19(46), 24(66), 26(60), 31(69), 34(84), 36(53), 39(65).

For the results presented below (see also Tables 3 and 4), we chose arithmetic mean values (M) and, as dispersion measures, the limits of the associated 95% confidence intervals.

Inferentially, we compared the scale scores obtained in each country group with the averaged scores of the respective other country cohorts. Because the data are not normally distributed and to avoid giving too much weight to individual outliers in the relatively small experimental groups, we tested nonparametrically using the sign test. Testing was two-sided. Exploratively, we also separately considered the scaling results of male (n = 67) and female (n = 91) participants. We tested differences in scale scores using both t-tests and the distribution-free Mann-Whitney U tests, both of which yield comparable results. The tests were

Table 4
Correlations (Pearson) between the PT measures of five cohorts from five countries and split-half reliability.

	CA	DE	JP	KR	TW	n	r	M_{Diff}
ALL	0.97	0.96	0.99	0.99	0.99	79	0.998	0.7
CA	–	0.97	0.95	0.96	0.95	5	0.923	3.5
DE		–	0.94	0.94	0.93	16	0.990	1.2
JP			–	0.97	0.98	15	0.991	1.8
KR				–	0.97	20	0.992	1.4
TW					–	25	0.995	1.2

Note. CA = Canada; DE = Germany; JP = Japan; KR = South Korea; TW = Taiwan; ALL = Measures obtained from all participants. Indicators from the split-half observation are n = no of persons in halved samples, r = correlation, M_{Diff} = mean difference (absolute) of scale values. All correlations are significant $p < 0.001$.

two-sided, except for the comparison of the physical exertion scales of test items T4, T5, and T6, which describe lifting and carrying situations. We assume there that these activities are less strenuous for assumed physically stronger men.

Overall, we assumed that people's experience of PT will not differ across cultures. Therefore, in order to allow possible evidence of cultural or methodological differences to become apparent when tested conservatively, we did not use Bonferroni corrections.

As a global measure and to illustrate the agreement between the scaling results of the individual cohorts, we correlated the respective mean values for each of the 32 presented activities examined (Pearson). An overview of the obtained correlations is given in the correlation matrix in Table 4. To gain a measure of the reliability of the scaling method for the relatively small subject groups, we also used the Pearson correlation to describe split-half reliability when participants were randomly assigned to each cohort. These results are also shown in Table 4.

4. Results

The training items that were conducted to ensure that people understood the scaling task were roughly estimated immediately after they were completed. The inspections showed that all participants were able to apply the CP procedure without any problems. The scaling results obtained in the orienting phase as listed in Table 2 do not differ very much between the cohorts and correspond to the expectation. The data of all cohorts also agree very well with the scaling results of 58 workers (average age about 52 years) of a sewage company (Müller, 2007; Müller and Kakarot, 2016). Only the physical exertion was scaled clearly differently by the workers, who are used to handle heavy weights, for items 4, 5 and 6 ($M_4 = 23.3$; $M_5 = 12.9$; $M_6 = 39.5$).

The scaling results of domestic activities are shown in Table 3 and Fig. 4 in ascending order of the overall results. According to the level of the respective required PT, the arithmetic means with the limen of 95% confidence intervals (CI) as obtained in the respective cohorts and over all participants are summarized. The numbers in the 2nd column (Pos) indicate the order in which the items were presented. Item-related mean values of the cohorts that deviate significantly from the overall mean values calculated for the participants of the four remaining cohorts in each case are printed in bold.

As can be seen from the corresponding correlations (see Table 4), there is high agreement of the PT measures between the examined cohorts. And we observe high consistency of data within cohorts. From its genesis, the CP method claims to enable meaningful descriptions even in small samples. To test this promise, we used a split-half approach by randomly dividing the cohorts into two subgroups each and comparing the subgroup results. Table 4 shows a good agreement of the scaling results ($r = 0.92$) already between the subgroups of the Canadian cohort, which were quite small with $n = 5$. The mean difference of the absolute scale scores of the subgroups ($M_{Diff} = 3.5$) is in the order of one third category. With sample size increases of three, four, and five times for Japan, Korea, and Taiwan, the correlations exceed $r = 0.99$, and the differences between scale scores approach one scale-unit.

4.1. Data obtained in the participating countries

Fig. 4 shows that the highest match is between the data sets from Japan (JP) and Taiwan (TW). The lowest of all possible matches is between the Taiwan and German (DE) datasets, although they are still highly significantly correlated ($r = 0.93$, $p < 0.001$).

Table 3 shows that these differences are greatest in the rather low PT range. The results indicate that German students experience greater PT in the rather low PT area than do students in Asia. Especially when listening to music, reading newspapers, watching the news, telling jokes, while shopping and expectantly unwrapping a gift, German students describe the situations as more strain demanding than do their counterparts in Japan, Korea, and Taiwan.

4.2. Gender differences

Both the PE-data from the orienting trials as well as the PT data from orienting and main trials show a remarkable agreement between the descriptions given by women and men as shown for the main trials in Fig. 5. For the orienting trials the average reported tension does not differ between the $n = 91$ women ($M = 22.4$ years) and $n = 67$ men ($M = 22.6$ years) that were included in the evaluation. However, the women

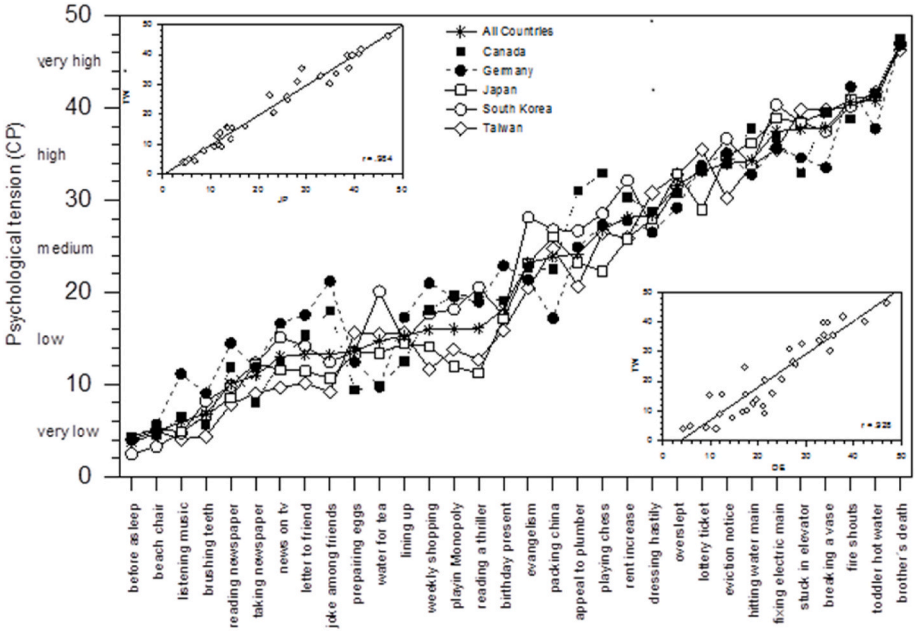

Fig. 4. PT Scale scores (M) on 32 different everyday situations obtained from cohorts in 5 countries.
Note. Almost perfect agreement between the data of Japanese (squares) and Taiwanese (diamond) students (see also inserted figure above). Lowest agreement in this study between the data of Taiwanese and German students (circles and inserted figure below). Items in ascending order of the overall results.

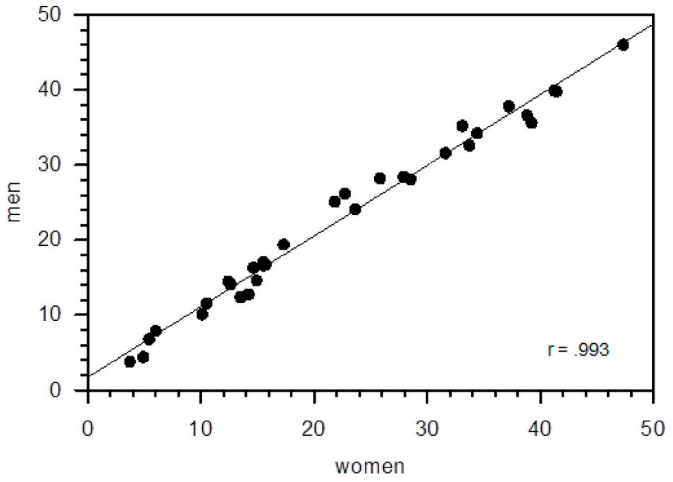

Fig. 5. Correlation (Pearson) of the mean scale scores of n = 67 male and n = 91 female participants.

reported higher exertion for items T4 "carrying a crate containing 12 bottles of water into the 3rd floor" (+2.9 CP-units, $U = 2237.5$, $p < 0.05$) and T5 "lifting a 5 kg-bag from the floor and putting it on a table" (+4,4 CP-units, $U = 1879$, $p < 0.01$).

Within the main trials are a few significant differences (Mann-Whitney-U, $n = 158$) between the ratings of men and women. Women associate the situations "getting stuck in an elevator" (+3.6 CP-units, $U = 2183.5$, $p < 0.002$), "breaking a precious vase", (+2.2 CP-units, $U = 2426.5$, $p < 0.028$) and "being informed by the police through a phone call of your brother's accidental death" (+1.3 CP-units, $U = 2217.5$, $p < 0.003$) with higher mental strain than men.

In contrast, men experience the social situations "conversation with a Jehovah's Witness" (+3.3 CP-units, $U = 2435$, $p < 0.03$), "appealing to the plumber" (+3.5 CP-units, $U = 2311.5$, $p < 0$ 0.009), but also "teethbrushing" (+1.9 CP-units, $U = 2311.5$, $p < 0.009$) and "listening to pleasant music while lying in bed" (+1.4 CP-units, $U = 2494.5$, $p < 0.05$) with higher levels of PT than women.

5. Discussion

Frankly, we were surprised by the remarkably good agreement between the data sets collected in the five countries. The fact that the results of the study are so harmonious is primarily due to the measurement technique we used and to the careful observance of the measurement regulations of the CP procedure by the test supervisors at the participating universities.

Perception psychologists repeatedly experience that the phenomena under investigation appear to dissolve and that the world, which in itself is stable when viewed objectively, disintegrates when it is critically questioned by scrutinizing participants. A quantitative description of unquestioned perceived characteristics becomes difficult or impossible if the situations to be described are repeatedly compared and placed into different contexts. Such judgment processes promote range effects if the range of stimuli experienced does not correspond to the range of scales provided. Frequency effects are more likely to occur if many closely related stimuli are to be described using only a few categories. Imagine a person is asked to describe the size of three definitely small lines, of which the line in the second position is slightly but noticeably larger than the first and the third line is slightly larger than the second. One way to communicate that the lines are of different length is to change the category from "very small" to "small" to "medium". The CP scale we used, however, allows a participant to define the length of all of the three lines as small as seen and to describe the differences between the lines by choosing suitable numbers without having to change the category. The aim of CP technology is to stabilize perception and measurement to ensure an unbiased quantifiable description of the perceived world. However, subsequent experiments will have to clarify whether the method meets these expectations. For these, the extent of the stimulus series and scales could be systematically varied. Data from experiments in which each subject scales only one single stimulus could be used to check whether the scaling procedure itself alters the object under investigation.

The observation that people do not differ substantially from each other in their assessment of psychological tension (Anspannung), regardless of where they were born, is not surprising for perceptual psychologists when they look objectively and measure carefully, although it does contrast with the expectations from the Sapir-Whorf hypothesis. In the latter case, it would be expected that not having a word for a psychological experience would preclude the observed degree of convergence between people from different cultures and different native languages. It would seem, however, that at least the experiences and worlds of experience we studied here arise from the same universal laws. These include the fact that the experiences one has in the course of life in different environments and cultures, which form and stabilize reference systems, are the basis for the perception and formation of metric sensation scales.

The results of the cross-country comparison give us indications that comparable activities might have a somewhat different meaning in different cultural contexts. Examples could be the situation "cracking a joke among friends", which the German participants associate with significantly higher PT (+8.9 CP-units) than do the participants of the other countries. Most Germans like jokes. If they are well presented, they enhance the social standing of the person cracking the joke. A German saying goes: "humor is a serious matter" (*Humor ist eine ernste Angelegenheit*).

The strikingly low level of PT during weekly shopping as reported by the Taiwanese students (- 6 scale-units) may be influenced by the specific shopping landscapes consisting of lots of small shops and restaurants around the residential area and by shopping habits. In Taiwan it's very common and considered to be pleasant to walk around in a market without necessarily buying anything for daily life. After school or working, Taiwanese drop in to a convenient store "to check" something new or to buy some refreshments. It's a kind of stress release in their daily life.

When people are transferred to other cultures, the repeated measurement of PT experienced in culturally typical situations could be an indicator of successful or unsuccessful integration. In a globalizing world, it is helpful and reassuring to know that we can communicate about stress and strain across countries. This offers, among many other options, the opportunity and is the basis for shaping the working and living conditions for all employees of a globally active company regardless of their origin and workplace.

However, a basic requirement for the quantitative description of the phenomenal world, and a challenge for the investigators, is that the eliciting scale "speaks the same language" as the respondent. It seems that we succeeded in finding the experience-adequate expressions for the participants of the five countries involved in the study.

Author's statement

Friedrich Müller: Conceptualization, Methodology, Investigation, Supervision, Writing – original draft preparation, Visualization. **Shuji Mori**: Conceptualiszation, Funding acquisition, Supervision, Writing – review & editing. **Yuko Sakaki**: Project administration and coordination Asia, Data curation. **Kwangoh Yi**: Investigation. **Sungbong Bae**: Investigation. **Yuka Tan**: Investigation. **Lawrence M. Ward**: Funding acquisition, Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Thanks to the Japan Society for the Promotion of Sciences for financial support of the presented study. LMW was supported by NSERC of Canada and SM was supported by a JSPS Grant-in-Aid for Challenging Research (Exploratory), No.20K20666. Some parts of this study were presented at the Fechner Day 2013 in Freiburg, Germany (Müller et al., 2013) and the International Congress of Psychology 2016) in Yokohama, Japan (Sakaki et al., 2016).

References

Bartenwerfer, H., 1963. Über Art und Bedeutung der Beziehung zwischen Pulsfrequenz und skalierter psychischer Anspannung. Zeitschrift für Experimentelle und Angewandte Psychologie 10, 455–470.

Bartenwerfer, H., 1969. Einige praktische Konsequenzen aus der Aktivierungstheorie. Zeitschrift für Experimentelle und Angewandte Psychologie 16 (2), 195–222.

Düker, H., 1975. Untersuchungen über die Ausbildung des Wollens. Bern, 1975.

Ellermeier, W., Westphal, W., Heidenfelder, M., 1991. *On the „absoluteness" of category* and magnitude scales of pain. Percept. Psychophys. 49, 159–166. https://doi.org/10.3758/BF03205035.

European Committee for Standardization, 2000. EN ISO 10075. *Ergonomic principles Related to Mental Work-Load - General Terms and Definitions*. https://doi.org/10.5100/jje.31.171.

Eysenck, H.J., Eysenck, M.W., 1987. Persönlichkeit und Individualität: Ein naturwissenschaftliches Paradigma. München-Weinheim (Psychologie Verlags Union).

Hellbrück, J., 1996. Category-subdivision scaling - a powerful tool in audiometry and noise assessment. In: Fastl, H., Kuwano, S., Schick, A. (Eds.), Recent Trends in Hearing Research. Festschrift for Seiichiro Namba. BIS, Oldenburg, pp. 317–336.

Heller, O., 1980. Orientierung innerhalb von phänomenalen Steigerungsreihen. In: Lauterbach, W., Sarris, V. (Eds.), Hrsg.) *Beiträge zur psychologischen Bezugssystemforschung*. Bern, Huber, p. 124ff.

Heller, O., 1981. Zur quantifizierung psychischer Anspannung. In: Tent, L. (Ed.), Erkennen, Wollen, Handeln. Festschrift für Heinrich Düker, pp. 213–225 (Göttingen, Hogrefe).

Heller, O., 1982. Theorie und Praxis des Verfahrens der Kategorienunterteilung (KU). In: Heller, O. (Ed.), Forschungsbericht 1981, pp. 1–15. Würzburg (Psychologisches Institut, Lehrstuhl für Allgemeine Psychologie).

Heller, O., 1985. Hörfeldaudiometrie mit dem Verfahren der Kategorienunterteilung (KU). Psychologische Beiträge 27, 478–493.

Heller, O., 1991. Orientated category scaling of loudness and speechaudiometric validation. In: Schick, A., Hellbrück, J., Weber, R. (Eds.), Contributions to Psychological Acoustics V. Results of the Fifth Oldenburg Symposium on Psychological Acoustics. Bibliotheks- und Informationssystem der Universität Oldenburg, S, pp. 135–159.

Heller, O., Boretzki, M., Fichtl, E., Knoblach, W., May, B., Nowak, Th, Stock, A., 1995. Entwicklung eines integrativen Verfahrens zur audiologischen Diagnostik, ausgerichtet auf die Prüfung und Bewertung technischer Hörhilfen für Hörgeschädigte. *Forschungsbericht an das das Bundesministerium für Forschung und Technologie Förderkennzeichen: 01 VJ8903*.

Kakarot, N., Müller, F., 2014. Assessment of physical strain in younger and older subjects using heart rate and scalings of perceived exertion. Ergonomics 57 (7), 1052–1067. https://doi.org/10.1080/00140139.2014.910613.

Lucy, J.A., 1992. Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis. Cambridge University Press, Cambridge, UK.

Mackworth, N.H., 1948. The breakdown of vigilance during prolonged visual search. Q. J. Exp. Psychol. 1, 6–21. https://doi.org/10.1080/17470214808416738.

Müller, F., 1981. Validierung der Skalierung psychischer Anspannung im Reaktionszeitexperiment. In: Tent, L., Hrsg (Eds.), Erkennen-Wollen-Handeln. Göttingen, Hogrefe, pp. 226–232.

Müller, F., 1988. Range effects in pitch scaling: a compromise between "rubber scale principle" and "true" (ecologically valid) series. In: Ross, Helen E. (Ed.), *Proceedings of the 4th Annual Meeting of the International Society for Psychophysics*. International Society for Psychophysics, pp. 103–104.

Müller, F., 1992. Scaling of psychological tension in reaction time experiments for the measurement of the efficiency of psychological performance with and without background noise. In: Borg, G. (Ed.), *Proceedings of the 8th Annual Meeting of the International Society for Psychophysics*. International Society for Psychophysics, Stockholm, pp. 147–152.

Müller, F., 2007. Assessment of age-correlated occupational strain as a basis for an age-adequate work-sequence organization. In: Mori, Shuji (Ed.), Perception and Performance in Real Complex Environments. Kijima Printing, Fukuoka, Japan, pp. 1–5.

Müller, F., Neely, G., Fichtl, E., 1995. Scaling of loudness, perceived exertion, and pain intensity: a comparison between the category partitioning (CP) and category ratio (CR-20) scaling procedures. In: Neely, G.W. (Ed.), *Category-Ratio Scaling of Sensory Magnitude in Comparison with Other Methods* (Study 3). Stockholm University, Dept. of Psychology.

Müller, F., Steinbach, G., 1997. Range effects in pitch scaling of instrumental sounds. In: Preis, A., Hornowski, T. (Eds.), Fechner Day 97, *Proceedings of the 13th Annual Meeting of the International Society for Psychophysics*. International Society for Psychophysics, Poznan, Poland, pp. 257–262.

Müller, F., Meschke, H., Fröschl, Th, 2003. Puls 300. Untersuchung der Beanspruchung von Personen, die sich während Vorbeifahrten des ICE 3 mit Geschwindigkeiten bis zu 300 km/h im Sicherheitsraum neben dem Gleis aufhalten. Bericht einer experimentellen Untersuchung an die DB Netz AG. DB GesundheitsService GmbH, Karlsruhe.

Müller, F., Kakarot, N., Mori, S., Sakaki, Y., Ward, L., 2013. Quantifying mental strain required to carry out everyday activities by German, Japanese and Canadian students. In: Wackermann, J., Wittmann, M., Skandies, W. (Eds.), *Fechner Day 2013 - Proceedings of the 29th Annual Meeting of the International Society for Psychophysics*. International Society for Psychophysics, Freiburg, Germany, p. 66.

Müller, F., Kakarot, N., 2016. Assessment of age-correlated occupational strain as a prerequisite for age-appropriate work organization. In: Wiencke, M., Cacace, M., Fischer, S. (Eds.), Healthy at Work. Springer, Heidelberg, pp. 367–377. https://doi.org/10.1007/978-3-319-32331-2_25.

Murayama, N., 1963. Examination of the relationship between goguryeo and Korean. Journal of the Academic Association of Koreanology in Japan 26, 189–198.

Neely, G.W., 1995. Category-ratio Scaling of Sensory Magnitude in Comparison with Other Methods. Department of Psychology, Stockholm University, Stockholm.

Oe, T., 1978. Korean and Japanese. Japanese Genealogy and History. *Iwanani Course Japanese 12*. Iwanami Shoten.

Ono, S., 1980. Outline Development of Japanese Genealogy. Modern Esprit Additional. Shibundo.

Parducci, A., 1963. The range-frequency compromise in judgment. Psychol. Monogr. 77 (2) whole No. 565. https://psycnet.apa.org/doi/10.1037/h0093829.

Pfanzagl, J., 1959. Die axiomatischen Grundlagen einer allgemeinen Theorie des Messens. Würzburg.

Rohmert, W., 1984. Das belastungs-beanspruchungs-konzept. Z. Arbeitswiss. (Neue Folge) 38 (4), 193–200.

Sakaki, Y., Müller, F., Mori, S., Yi, K., Bae, S., Tan, Y., Ward, L.M., 2016. Evaluation of mental strain going with everyday activities. A comparison of different cultural backgrounds. In: 31st International Congress of Psychology. Yokohama, Japan.

Semjonova, V., 2012. Extraversion und Vigilanz. Masterthesis. Universität Lüneburg.

Tanaka, H., Iemura, M., Igarashi, Y., Kuramata, H., Nakamura, T., Higuchi, T., 1975. Introduction to Linguistics. Taishukan Shoten.

Tent, L., 1988. Heinrich Dükers handlungstheorie. Arch. Psychol. 140 (3), 169–182.

Teghtsoonian, R., 1973. Range effects in psychophysical scaling and a revision of Stevens' law. Am. J. Psychol. 86, 3–27. https://doi.org/10.2307/1421845.

Umeda, H., 1982. Korean and Japanese - problems in contrastive studies. Journal of Japanese Language Teaching 48, 31–42.

Witte, W., 1960a. Experimentelle untersuchungen von Bezugssystemen. Psychologische Beiträge 4, 218–252.

Witte, W., 1960b. Über phänomenskalen. Psychologische Beiträge 4, 645–672.

Yu, B., Sugimura, T., 1993. Comparisons of Category Typicality in China and Japan, vol. 42. Bulletin of Nara University of Education. (Culture & Society), pp. 133–142, 1.