Development, Validation, and Reliability Testing of the Brief Instrument to Assess Workers' Productivity during a Working Day (IAPT)

Fábio Sprada de Menezes[1,2]
[1]*Federal University of Technology – Paraná (UTFPR), Production Engineering Department, Ponta Grossa, Brazil*
[2]*Estácio University Center of Santa Catarina, São José, Brazil*

Antonio Augusto de Paula Xavier[3]
[3]*Federal University of Technology – Paraná (UTFPR), Production Engineering Department, Ponta Grossa, Brazil*

Abstract

Purpose – The aim of this study was to develop, validate, and test the clarity and reliability of the Brief Instrument to Assess Workers' Productivity during a Working Day.

Design/methodology/approach – The content of the instrument was chosen using research containing other valid instruments and after this the construct was developed. Relevance and clarity validations were conducted with experts using Likert scales (from 0 to 10), convergent validity was performed using the Health and Productivity Questionnaire (HPQ) and Health & Labor Questionnaire (HLQ) instruments, and reliability measures were carried out using the Split Half Test and Cronbach's Alpha coefficient.

Findings – The instrument proved to be clear and relevant with an average of 9.11 ± 0.93 in the relevance test and 9.23 ± 0.75 in the clarity test. Regarding convergent validity, the instrument showed a high correlation with the HPQ ($r^2 = 0.86$) and the HLQ ($r^2 = 0.82$). The reliability results were $r^2 = 0.78$ in the Split Half Test and a Cronbach's Alpha coefficient of $\alpha = 0.91$ for the Management variables and $\alpha = 0.80$ for the Physical and Mental Variables.

Originality/value – The proposed instrument was shown to have an adequate content and construct, in addition to converging results with other recognized instruments, and it had very high levels of reliability. All these factors define it as a good tool for research regarding productivity in companies.

Keywords – Productivity, Workers, Validity, Reliability

Review of Business Management

1 Introduction

Workers' productivity is a widely studied subject related to management and human resources within companies. Employees with low productivity are associated with financial losses and higher costs in order to compensate for the deficit resulting from low performance, all of which must be accounted for during financial planning (Krol & Brouwer, 2014). In the United States, company losses involving the costs of low productivity are estimated at US$ 260 billion annually (Mitchell & Bates, 2011).

A worker's performance may decline due to two main reasons: absenteeism and presenteeism. Absenteeism is measured by the number of absences a worker presents in a specific period of time; it is usually caused by infectious diseases or recurrent injuries that affect overall health (which may or may not be work related). Presenteeism is the type of productivity decline that is not related to absences, but instead related to distractions, stress, fatigue, and a series of physical and mental conditions that result in lost efficiency during working hours (Schultz, Chen, & Edington, 2009).

The number of studies on presenteeism is considerably low in comparison to those on absenteeism (Stewart, Ricci, Chee, & Morganstein, 2003) and the productivity losses associated with it are difficult to measure (Stang, Cady, Batenhorst, & Hoffman, 2001).

Whereas with absenteeism it is possible to calculate lost productivity by counting the number of days a worker is absent, with presenteeism this assessment is much more difficult, for it is associated with the physical and/or psychological commitment presented by employees (Despiegel, Danchenko, Francois, Lensberg, & Drummond, 2012). These losses represent an estimated 77% of the total losses associated with productivity decline, against the 23% loss associated with absenteeism (Callen, Lindley, & Niederhauser, 2013).

It is a given that stress and physical tiredness influence performance during the workday. As the day develops, together all the effort, unfinished tasks, and time passed seem to lead to a decline in productivity and in the capacity to complete simple work-related assignments (Despiegel et al., 2012; Lamontagne, Keegel, Louie, Ostry, & Landsbergis, 2007).

However, assessing workers' productivity decline during the workday is a complex challenge. In some cases, when productivity can be measured by a number of completed tasks (as in factory production lines or call centers) it is fairly simple. However, for jobs that require the execution of many different tasks, like office work or customer service, assessing productivity becomes problematic (Burton, Pransky, Conti, Chen, & Edington, 2004).

For this reason, more and more self-reported instruments have been developed and validated in an attempt to better assess productivity. Even though they are based only on the workers' own perceptions regarding their work productivity at a specific moment, self-reported instruments are currently the best option for these types of work conditions.

There are many self-reported instruments, but they present some limitations. The challenge lies in how the information is collected. Most instruments assess overall productivity, dealing with absenteeism and presenteeism at the same time; others are rather extensive, taking way too long to be completed and causing lengthy interruptions or even taking days to recall (and therefore they cannot be applied entirely on a single day) (Despiegel et al., 2012; Mattke, Balakrishnan, Bergamo, & Newberry, 2007).

In addition to all of the reasons mentioned above, no self-reported instruments capable of assessing productivity fluctuations within one entire workday were found, which certainly represents a gap to be filled in the current literature.

In this context, the purpose of this study was to develop, validate, and test the reliability of a fast and easy self-reported instrument capable of assessing workers' productivity during a workday.

233

Rev. Bras. Gest. Neg. São Paulo v.20 n.2 apr-jun. 2018 p.232-247

2 Theoretical Basis

2.1 Work Productivity

The productivity of a task can be defined as the final product of three very important variables: time spent executing the task, the quality of the final product, and the cost of the task, as shown in Figure 1 (Ulubeyli, Kazaz, & Er, 2014)
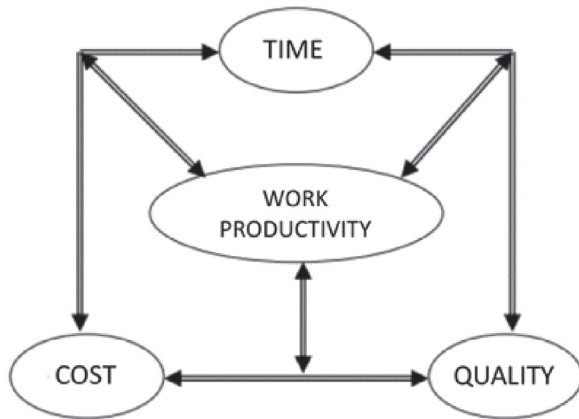

Figure 1. The basic elements of Work productivity.

Source: Adapted from "Planning Engineers' Estimates on Labor Productivity: Theory and Practice", S. Ulubeyli, A. Kazaz, and B. Er,, 2014, *Procedia - Social And Behavioral Sciences, 119.*

The concept of labor productivity is present in the current literature and it has been used by workers, companies, and countries to measure and keep track of their own performance. For a long time, productivity was measured as the ratio between production and the number of workers. This approach was a tormenting way to stimulate employees' productivity. As time passed, other ways of measuring productivity were developed, relating productivity to the use of resources such as energy, raw material, and inputs,

among other things (King, Lima, & Costa, 2014).

Another simplified way of defining productivity calculated it as the ratio between the tasks taken on and the time devoted to the work. Therefore, the less time a job takes to be delivered successfully, the more productive it becomes and vice versa (Jackson & Victor, 2011). Productivity can also be defined as the overall performance of a group of workers, which reflects how efficient the group is (Stang et al., 2001).

It is known that human capital, manifested by the experience and knowledge of a company's employees, is the most important factor for a company to be considered productive (Chowdhury, Schulz, Milner, & Van De Voort, 2014).

Companies look for more productive workers and these are usually more recognized and valued, frequently receiving the best salaries. This situation is often motivated by corporate policies that offer bonuses to more productive employees (Englmaier, Strasser, & Winter, 2011). In addition, it is known that more productive workers are also promoted faster. Bosses are usually 1.75 times more productive than normal workers (Lazear, Shaw, & Stanton, 2014).

There are several factors that can influence job productivity. Factors such as motivation, health problems, workers' mental health, management style, circadian rhythm, work place temperature, monotonous tasks, and the duration of breaks or of uninterrupted work are often cited as productivity moderators (Krol, Brouwer, & Rutten, 2013; Kuhn, 2001; Sadosky, DiBonaventura, Cappelleri, Ebata, & Fujii, 2015; Sahu, Sett, & Kjellstrom, 2013; Wahlstrom, Hagberg, Johnson, Svensson, & Rempel, 2002).
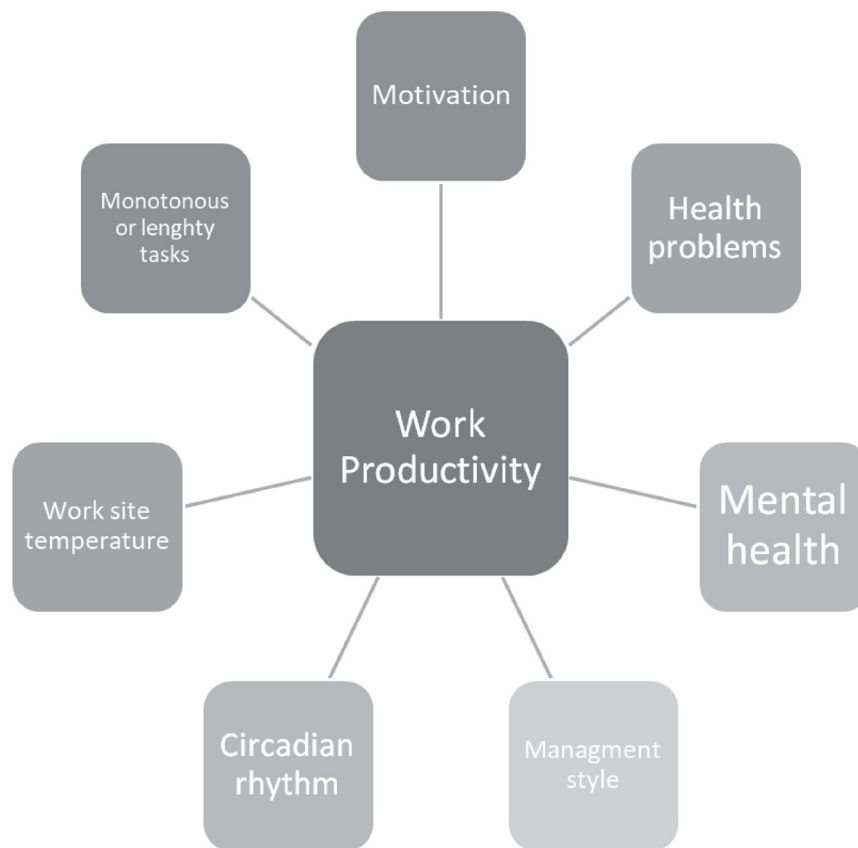
234

Rev. Bras. Gest. Neg. São Paulo v.20 n.2 apr-jun. 2018 p.232-247

Figure 2. Factors that influence productivity found in the literature

2.2 Productivity Assessment Instruments

There is a need, on the part of the managers, to consider how productivity varies during a working day and the factors that cause these variations. Knowing these variations allows for strategies to be devised in order to avoid performance declines. However, quantifying these variations is quite complex.

When dealing with tasks like equipment assembly or product delivery, this assessment is easier, as variations in productivity can be measured by the number of tasks completed. On the other hand, in occupations that involve bureaucratic activities or customer service, such an assessment becomes troublesome (Burton et al., 2004).

For these specific cases, where it is difficult to identify productivity variations in an objective way, tools have been created that identify them in a self-reported way. These self-reported tools assist managers in the diagnosis of their employees' productivity.

Several instruments with this purpose exist in the literature. In order to acknowledge them, with their advantages and disadvantages, a search of the main databases was conducted with the purpose of getting to know the latest instruments that evaluate productivity. This search eventually stimulated the creation of our instrument and its strategy is best described in the Method section of this study.

Considering the purpose of the study, which is to examine productivity specifically during the workday and the variations in productivity caused by presenteeism, the following instruments were found: the Health and Productivity Questionnaire (HPQ) (Kessler et al., 2004), the Health and Labor Questionnaire (HLQ) (Hakkaart-van Roijen & Essink-Bot, 2000), the Work Productivity and Activity Impairment Questionnaire (WPAI) (Reilly,

235

Rev. Bras. Gest. Neg. São Paulo v.20 n.2 apr-jun. 2018 p.232-247

Zbrozek, Dukes, 1993), the Work Limitations Questionnaire (WLQ) (Lerner et al., 2001), the Stanford Presenteeism Scale (SPS) (Frauendorf, Medeiros, Pinheiro, & Ciconelli, 2014), the Work and Health Interview (WHI) (Stewart, Ricci, Leotta, & Chee, 2004), and the Sheehan Disability Scale (SDS) (Sheehan & Sheehan, 2008).

After careful analysis of these tools, some limitations were observed. Instruments such as the SDS and the SPS did not present any data in their original articles that confirmed that a content validation and clarity check had been performed. The WHI did not have data from a reliability analysis, a fundamental step in obtaining data from a question and answer tool. Another instrument, the WPAI, focused only on illness and its relationship with declining productivity. Other tests like the HPQ, WLQ, and HLQ were time-consuming, which would make it impossible to collect an entire workday's worth of data, as they would hinder work progress.

However, the main shortcoming found in the instruments above was the number of questions that involved absenteeism and the recall time between evaluation and reevaluation being at least one week. These conditions would not allow the observation of productivity fluctuations during the workday. This revealed the need to develop an instrument with a short recall time (2 hours) that was quick to complete and that could be applied more than once during the workday.

3 Method

3.1 Content Development

The process of developing the instrument, as already mentioned, arose from the need to measure workers' productivity during the day. It is part of the project titled "Statistical Approach Subjective Productivity at Work, a Perspective from Workers' Individual Psychophysiological Conditions", duly approved by the Human Research Ethics Committee of the Federal University of Technology - Paraná (UTFPR)

under approval number CAAE 52897315.5. 0000.5547.

Initially, research was carried out on the literature with emphasis on the discovery and exploration of similar instruments in a search for questions that linked to subjective variables and behaviors that could indicate workers' productivity levels. In addition, the format and punctuation of these instruments were observed for construct development purposes.

For this, a literature search was carried out for articles published between 2000 and 2015 and indexed in the databases: Web of Knowledge, Pubmed, Bireme, EBSCO Host, Science Direct, and Scopus. The strategy used isolated, cross, and truncation searches for descriptors used by the authors in the titles or abstracts, adopting the Boolean expression AND. The descriptors were: Productivity; Job; Presenteeism; Questionnaires; Instruments. The descriptors were searched for in Brazilian Portuguese and English.

A total of 522 published articles were initially found and compiled by titles and abstracts. After reading the full articles and observing their relevance, a more careful analysis was made with the main problem in mind, observing the similarities and needs of this research.

At the end of the search phase, 14 studies that contained important concepts and developed and tested tools for research on productivity at work were selected and used as a basis for the questionnaire model and the instrument validation process.

From that point on, the 10 (ten) questions (Table 1) were conceived that served the purpose of the instrument and were adaptable to the fact that it must be applied a few times during a workday.

According to Stewart et al. (2003), lost productivity is often related to a lack of concentration during the execution of activities, to the repeated execution of the same activity (loss of efficiency), and to fatigue. Further investigation of these conditions motivated the selection of two of the questions (1 and 2).

Feeling motivated and fit for work, along with a self-perception of productivity, leads to better performance and greater satisfaction with the work done. These statements motivated the selection of questions 3, 4, and 10 (Gagné & Deci, 2005). Feeling confident to perform a function is also a condition that is always related to productive professionals and it motivated the selection of question 5 (Folkard & Tucker, 2003).

Question 6 is associated with work-related anger and irritation. It is known that 47% of lost productivity at work is associated with mental conditions and that about 67% of complaints associated with work-related mental stress are associated with feelings of anger and irritation (Gates, Gillespie, & Succop, 2011; Goetzel, Ozminkowski, & Long, 2003).

Question 9 was conceived based on studies which show that besides mental conditions, physical conditions also affect productivity (Lindegard, Larsman, Hadzibajramovic, & Ahlborg, 2014). According to Goetzel, Ozminkowski and Long (2003), pain and general symptoms account for 29% of productivity loss at work.

It is understood that vigor and mental resilience when facing work difficulties are also fundamental conditions for maintaining work engagement, a variable which, according to the authors, is the most important one for ensuring good productivity. Questions 7 and 8 were conceived based on this concept (Munir et al., 2015).

After defining the questions and in order to facilitate later analyses, they were divided into two dimensions: one called "Managerial Variables (MV)", which contemplates five questions that involve perceived satisfaction with the work performed, aptitude and confidence in decision making, and the workers' level of concentration and efficiency; and another dimension called "Physical and Mental Variables (PMV)", which refers to questions that examine variations in mood, clinical symptoms, and workers' levels of physical and mental fatigue.

The questions were randomly distributed and the "positive" or "negative" responses were alternated in order to make the instrument more reliable, with questions 1, 3, 4, 5, and 10 referring to the MV dimension and questions 2, 6, 7, 8, and 9 referring to the PMV dimension.

Fluctuations in productivity during the workday are subjective. In order to better capture and simplify future analysis of these variations, the workday was divided into periods of two hours each. Therefore, the researched worker should report his/her experiences regarding his/her work in the last 2 (two) hours and this action was to be repeated as many times as necessary until the end of the workday.

Table 1

Questions chosen to compose the instrument

How concentrated and efficient have I felt in the last two hours?
How tired or sleepy have I felt in the last two hours?
How productive have I felt in the last two hours?
How able to make work-related decisions have I felt in the last two hours?
How confident in my work-related decisions have I been in the last two hours?
How annoyed or upset during work have I been in the last two hours?
How difficult has it been to get work done in the last 2 hours?
How excited to work have I been in the last 2 hours?
How affected by physical symptoms (pain, dizziness, etc.) have I been in the last 2 hours?
How satisfied with my work performance have I been in the last 2 hours?

After these definitions, the development of the instrument format and scoring form began.

3.2 Format Development

With the instruments found in the first phase of the research still in mind and understanding the time sensitivity in order not to interfere much in the research subject's day, an instrument was created that was easy to understand and complete.

We opted for a table that had the 10 questions of the instrument in the first column and a progressive measure of the subjective perception on the first line, based on the principles of Likert. For each question, the terms Nothing, Little, Regular, Very, and Totally were used. The Likert model was chosen because it is not only consistent with the research goals, but it is practical and it follows the models used internationally, some of which have already been mentioned in this study.

So, the ten questions followed by the 5 columns were sequentially placed to mark the self-reported perception in relation to each question and in relation to the last two hours of work. In the instrument header, there are instructions for the respondent to answer it by marking only one of the fields per question and to leave no questions blank, ensuring a maximum return from the instrument. The complete instrument can be seen in Figure 3.

Brief Instrument to Assess Workers' Productivity During a Working Day

You are being asked to answer questions about your productivity performance during your workday. This instrument is individual and self-applicable and it refers to your performance in the **LAST TWO HOURS OF WORK**. At the end of the evaluation you will be asked your age and other data concerning yourself, these will help us better understand your answers.

You must mark with an "X" the field that most expresses your perception during the evaluation. Remember that all questions must be answered. **THERE ARE NO RIGHT OR WRONG ALTERNATIVES**, the instrument is subjective, so feel comfortable to answer the questions. We just want to know your personal opinion.

Feel free to withdraw from the survey if you feel embarrassed or uncomfortable.

Thank you for your help!

QUESTIONS	Nothing	Little	Regular	Very	Totally
How concentrated and efficient have I felt in the last two hours?					
How tired or sleepy have I felt in the last two hours?					
How productive have I felt in the last two hours?					
How able to make work-related decisions have I felt in the last two hours?					
How confident in my work-related decisions have I been in the last two hours?					
How annoyed or upset during work have I been in the last two hours?					
How difficult has it been to get work done in the last 2 hours?					
How excited to work have I been in the last 2 hours?					
How affected by physical symptoms (pain, dizziness, etc.) have I been in the last 2 hours?					
How satisfied with my work performance have I been in the last 2 hours?					

Figure 3. Complete and final version of the instrument

3.3 Scoring

The Likert scale was used to measure the responses and scores from 0 to 4 were assigned to each item. As some questions had "positive" connotations for productivity and others "negative" connotations, the adjectives and punctuation were alternated to avoid any biases.

The sum of the 10 questions enables a final score where 0 (zero) is the smallest possible value and 40 (forty) is the highest. The full table, with a detailed score for each question, is given in Appendix 1 at the end of this article.

At the end, to facilitate analysis, a Workers' Productivity Percentage is proposed. In order to obtain it, the following equation must be used:

$$\text{Productivity Percentage (\%)} = (\text{Final Score}/40) \times 100$$

3.4 Validation Process

From a general point of view, validity refers to the degree to which an instrument accurately measures the variable to be measured. Brewer and Hunter (2006) point out that the validity of an instrument is judged by its ability to perform its explanatory role, and its concept aims to bring together several aspects of validity. In order to organize the comparisons each validation step must be performed. The authors indicate that the validation process involves three important steps, as further explained below (Brewer & Hunter, 2006).

3.4.1 *Specialists Committee Validation*

For this stage, ten notoriously qualified specialists from different areas of labor studies were selected to judge the validity of the instrument. They included three professionals from the production engineering area, two from workers' health, two from occupational psychology, two from management and human resources, and one from personnel management, all of whom hold a Ph.D. or Masters and are professors in their respective areas. Their participation was by invitation and voluntary. After receiving the instrument, they could return it at their own convenience.

The experts were asked to analyze the clarity and relevance of each question separately. For clarity, the orientation given was to observe how understandable the question was and whether it expressed exactly the concept intended to be measured. As for relevance, this refers to how relevant the items are, whether they reflect the associated concepts, and whether the questions are appropriate to achieve the goals of the instrument (Alexandre & Coluci, 2011).

In order to validate the instrument, a simple document was created with an explanatory heading and consisting of a 0 to 10 point qualitative-quantitative Likert-like scale after each of the questions.

Each evaluator should indicate, on the numerical scale, the level of validity of each question. Following the scale, there was also a specific field for comments on the wording of the questions and further suggestions.

3.4.2 *Convergent Validity*

This process is associated with comparing the results obtained in our construct with the results from other already well established and validated constructs, to verify if all of them measure the same phenomenon.

As no similar instruments were found, the presenteeism dimensions from two of the previously selected productivity assessment instruments were adapted to serve as a comparison. The selected instruments were the Health and Productivity Questionnaire (HPQ), and the Health and Labor Questionnaire (HLQ). For the HPQ, the question regarding performance at work is B-15. It consists of a 0 to 10 progressive Likert scale, which asks the following:

239

Rev. Bras. Gest. Neg. São Paulo v.20 n.2 apr-jun. 2018 p.232-247

B-15 - Using the same 0 to 10 scale, how would you rate your overall job performance on the days you worked during the past 4 weeks (28 days)?

Worst performance					Top performance					
0	1	2	3	4	5	6	7	8	9	10

In order to adapt to this study's needs, the sentence "on the days you worked during the past 4 weeks (28 days)" was modified to "during the time you were evaluated". The productivity score is obtained by multiplying the score chosen by the worker by 10 (ten).

From the HLQ, questions 5 to 10 were used, which are intended to detect productivity problems at work due to health problems. The wording and format of the questions are as follows:

I did go to work but, as a result of health problems…:

	(Almost) Never	Sometimes	Often	(Almost)Always
5- I had a problem concentrating				
6- I had to work at a slower pace				
7- I had to seclude myself				
8- I found decision-making more difficult				
9- I had to put off some of my work				
10- I had to let others take over some of my work				

In order to adapt to this study's needs, the sentence "I did go to work but, as a result of health problems" was modified to "during the evaluated period, I".

The final score for this module in the instrument is obtained from the sum of the score for each question. For questions marked "never" the score is 1; for "sometimes" it is 2; for "Often" it is 3; and for "Always" it is 4 points. The maximum score is 24 points and the minimum is 6 points.

The convergent validation for this 10-question scale was obtained according to the concept developed by Hair, Anderson, Tatham, and Black (1998). A test was performed with 100 (one hundred) office workers, where the subjects completed all three instruments sequentially. The subjects were asked to maintain the same perceptions in all of the three questionnaires. At the end, the Pearson's correlation coefficient was applied to the data in order to identify linear relationships between the three instruments.

3.5 Reliability Measure

For a data collection instrument to be reliable it needs to be coherent and show consistency in its results. A reliable instrument generates reliable measurements and stable results (Martins, 2006).

To assess reliability, two tests were chosen: Split-Half Reliability and Cronbach's Alpha Reliability Coefficient.

The split-half test is done by splitting the questions of an instrument into two halves with similar characteristics in terms of the set of questions, the degree of difficulty, and the content characteristics. Both halves are then given to one group at the same time. If there is a strong positive correlation between the results of the two halves, the instrument is considered reliable.

The Cronbach's alpha coefficient was used to measure the internal consistency between the two dimensions of the instrument. This index is

able to verify the homogeneity of questions that seek to measure the same construct. It considers the variance between individuals as well as the variance attributable to the interactions between individuals and items. This estimate is affected by the number of variables and the intercorrelation between variables and of the instrument.

Reliability was then tested using the same 100 (one hundred) subjects whose tests had measured convergent validity. For the split-half test, the questions were randomly divided. Each half had five questions. Half A consisted of questions 1, 2, 3, 4, and 6 (two from the PMV dimension and three from the MV dimension). Half B consisted of the leftover questions (three from the PMV dimension and two from the MV dimension).

The Cronbach's alpha was calculated without mixing the two dimensions, given their different purposes.

3.6 Statistical Analysis

The data were initially treated and displayed using descriptive statistics (mean, standard deviation, and coefficient of variation). For some analyses where correlation measurements were required, the Kolmogorov-Smirnov test was performed and the analyzed data came out normal. Therefore, the Pearson's correlation coefficient was chosen for this function. The statistical package IBM SPSS™ 23 was used for the analysis.

4 Results

The results obtained from the development process mentioned in the methodology will be presented separately in order to facilitate visualization and understanding.
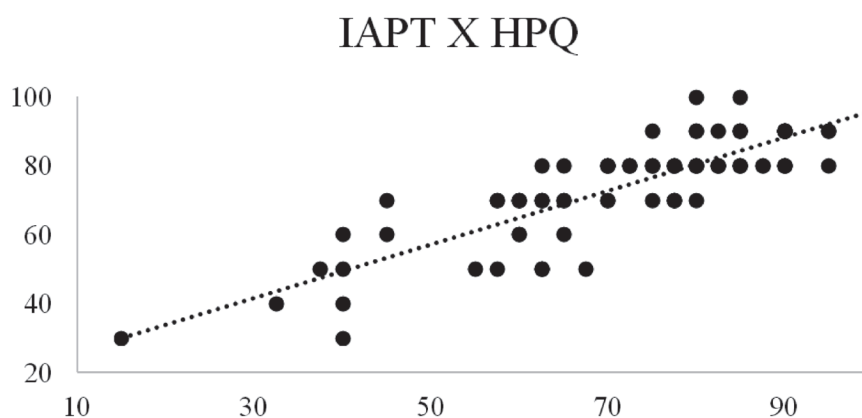
4.1 Specialists Committee Validation

The results for relevance were satisfactory with a low standard deviation and coefficient of variation for all of the questions. The highest mean was obtained for questions 1, 3, 5, and 10. The lowest mean was obtained for question 4. The instrument's final mean regarding relevance was 9.11 ± 0.93 (CV = 10.21%).

For the clarity test, satisfactory values were achieved once again. The highest mean was obtained for questions 1, 6, 9, and 10. The lowest mean was obtained for questions 2 and 7. The instrument's final mean was 9.23 ± 0.75 (CV = 8.12%).
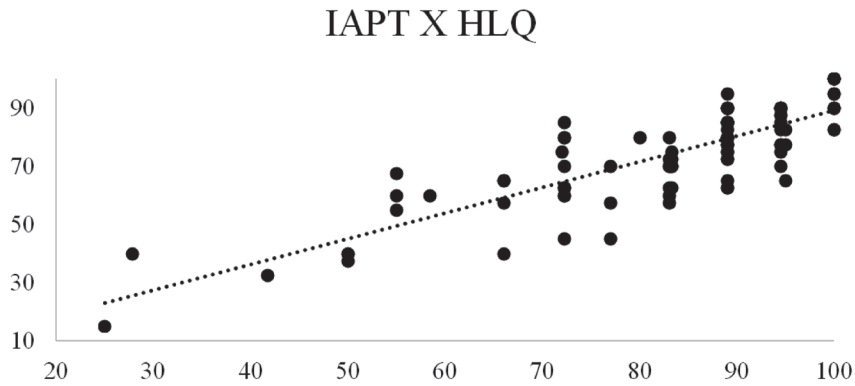
4.2 Convergent Validity

The convergent validity between the HPQ and this instrument is presented in Graph 1 for a better visualization and understanding of the correlation curve. The Pearson's correlation coefficient derived from this analysis was $r^2 = 0.86$ ($p \leq 0.05$), meaning a strong positive correlation between the results obtained in both instruments.

IAPT X HPQ

Graph 1 – Correlation between the proposed instrument and the HPQ

241

Rev. Bras. Gest. Neg. São Paulo v.20 n.2 apr-jun. 2018 p.232-247

After checking correlation with the HPQ, this was also tested with the HLQ and the results are presented in Graph 2. The Pearson's correlation coefficient derived from this analysis was r^2 = 0.82 (p≤0.05), which again shows a strong correlation between the results obtained in both instruments.
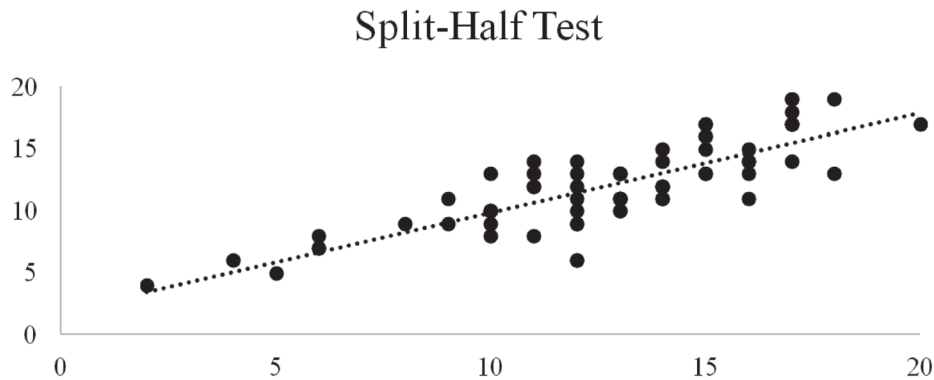
IAPT X HLQ

Graph 2 – Correlation between the proposed instrument and the HLQ

4.3 Reliability Measures

The instrument's reliability was tested and it was found that in the Split-Half test the correlation index obtained was r^2 = 0.78 (Graph 3). The Cronbach's alpha coefficient for the dimension titled Managerial Variable (MV) was α = 0.91. For the other dimension, Physical and Mental Variables (PMV), the index was α = 0.80.

Split-Half Test

Graph 3 – Results for the Split-Half reliability test

5 Discussion

When designing a measuring instrument, it is important to define what is being measured and how the measurement is going to be carried out. It is of fundamental importance that all objectives are established and that they are linked to the concepts one wishes to address. In addition, characterizing the target population is also crucial since it justifies the relevance of developing a specific instrument for a specific situation (Coluci, Alexandre, & Milani, 2015). The aim of this study was to test if the proposed instrument was adequate to gauge workers' self-reported productivity.

For the proposed instrument the main goals were to create a format based on few questions, easy comprehension, and that was

quick to fill out, so it would not cause major interruptions during the workday. According to Czerwinski, Horvitz, and Wilhite (2004), workday interruptions and disturbances caused by external agents such as music, phone calls, or interpersonal contact are one of the main causes of drops in productivity and lack of concentration while performing tasks. Hence the few simple questions and Likert scale, which are easy to understand and complete with little work routine disturbance.

Always with the aim of developing a practical and relevant instrument, experts in the labor field were asked about the instrument's clarity and relevance. The average scores on all questions were satisfactory and no questions from the previously developed instrument had to be modified for the final version of the instrument after the experts' assessment.

In addition to relevance and clarity, every measure must meet two minimum requirements: validity and reliability. Valid measures are those which accurately represent the phenomenon to be measured. Reliable measures are consistent in time and space and can be repeated by other researchers (Alexandre & Coluci, 2011; Czerwinski et al., 2004; Martins, 2006; Salmond, 2008).

The convergent validity showed a strong positive correlation between the values obtained in the adapted questions originating from the HPQ and HLQ when compared with our instrument. These data show that the instrument is able to measure what it is supposed to.

As for reliability, the Split-Half test and Cronbach's alpha index are well established ways of analyzing reliability. The former uses a correlation index, so the stronger the correlation, the more reliable the instrument (Fan & Thompson, 2001). For the latter, alpha values above 0.7 are satisfactory (Adamson & Prion, 2013; Aguiar, Fonseca, & Valente, 2010). For both tests, Split-Half and Cronbach's alpha, very satisfactory, higher than recommended values were observed regarding the proposed instrument. These results indicate that the instrument has

good internal consistency, it is easy to apply, and it can be reproduced.

6 Conclusion

At the end of this process and after careful examination, the brief instrument to assess workers' productivity was developed. This instrument was proven to be clear, easy to complete, and with good validity and reliability. As a consequence, it shows potential to be a contributing tool for studying and better understanding labor productivity, considering that it records fluctuations in this during the workday.

It is understood that self-reported measures do not have the same reliability as a direct productivity measure. However, the instrument can be applied in companies or services where productivity fluctuations cannot be measured by calculating the number of completed tasks within a certain period of time. The fact that this instrument is simple, clear, and brief enables it to be used at different times during a workday.

One limitation of this study, and possibly of the instrument as well, is that its validation process was performed using subjects with very homogeneous work characteristics. Therefore, more research needs to be done in order to validate this instrument in different working conditions that also enable the productivity measured to be associated with other variables that can influence productivity numbers such as shifts, physiological variables, pain, occupational diseases, the subjects' psychological state, and their mental load at work.

References

Adamson, K. A., & Prion, S. (2013). Reliability: Measuring Internal Consistency Using Cronbach's α. *Clinical Simulation in Nursing, 9*(5), 179 - 180.

Aguiar, O.B., Fonseca, M.J.M., & Valente, J.G. (2010). Confiabilidade (teste-reteste) da escala sueca do Questionário Demanda-Controle entre Trabalhadores de Restaurantes Industriais do Estado do Rio de Janeiro. *Rev Bras Epidemiol, 13*(2), 212-222.

Alexandre, N. M. C. & Coluci, M. Z. O. (2011). Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Ciência & Saúde Coletiva, 16*(7), 3061-3068.

Brewer, J., & Hunter, A. (2006). *Foundations of multimethod research.* Thousand Oaks.

Burton, W. N., Pransky, G., Conti, D. J., Chen, C. Y., & Edington, D. W. (2004). The association of medical conditions and presenteeism. *J Occup Environ Med, 46*(6 Suppl), S38-45.

Callen, B. L., Lindley, L. C., & Niederhauser, V. P. (2013). Health risk factors associated with presenteeism in the workplace. *J Occup Environ Med, 55*(11), 1312-1317.

Chowdhury, S., Schulz, E., Milner, M., & Van De Voort, D. (2014). Core employee based human capital and revenue productivity in small firms: An empirical investigation. *Journal of Business Research, 67*(11), 2473-2479.

Coluci, M. Z. O., Alexandre, N. M. C., & Milani, D. (2015). Construção de instrumentos de medida na área da saúde. *Ciência & Saúde Coletiva, 20*(3).

Czerwinski, M., Horvitz, E., & Wilhite, S. (2004). *A Diary Study of Task Switching and Interruptions.* Paper presented at the ACM Conference on Human Factors in Computing Systems, Viena.

Despiegel, N., Danchenko, N., Francois, C., Lensberg, B., & Drummond, M. F. (2012). The use and performance of productivity scales to evaluate presenteeism in mood disorders. *Value Health, 15*(8), 1148-1161. Englmaier, F., Strasser, S., & Winter, J. (2011). Worker characteristics and wage differentials: Evidence from a gift-exchange experiment. *Behavioural Economics, 3637*, 1-50.

Fan, X., & Thompson, B. (2001). Confidence intervals for effect sizes confidence intervals about score reliability coefficients, please: An EPM guidelines editorial. *Educational and Psychological Measurement, 61*(4), 517-531.

Folkard, S., & Tucker, P. (2003). Shift work, safety, and productivity. *Occup Med (Lond), 53*(2), 95-101.

Frauendorf, R., de Medeiros Pinheiro, M., & Ciconelli, R. M. (2014). Translation into Brazilian Portuguese, cross-cultural adaptation and validation of the Stanford presenteeism scale-6 and work instability scale for ankylosing spondylitis. *Clin Rheumatol, 33*(12), 1751-1757.

Gagné, M., & Deci, E.L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior, 26*, 331-362.

Gates, D. M., Gillespie, G. L., & Succop, P. (2011). Violence against nurses and its impact on stress and productivity. *Nurs Econ, 29*(2), 59-66, quiz 67.

Goetzel, R. Z., Ozminkowski, R. J., & Long, S. R. (2003). Development and reliability analysis of the Work Productivity Short Inventory (WPSI) instrument measuring employee health and productivity. *J Occup Environ Med, 45*(7), 743-762.

Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis* (5ed ed.). New Jersey: Prentice Hall.

Hakkaart-van Roijen, L., & Essink-Bot, M.L. (2000). Manual: the Health and Labour Questionnaire.

Jackson, P., & Victor, T. (2011). Productivity and work in the 'green economy' Some theoretical reflections and empirical tests. *Environmental Innovation and Societal Transitions, 1*, 101-108.

Kessler, R. C., Ames, M., Hymel, P. A., Loeppke, R., McKenas, D. K., Richling, D. E., Ustun, T. B. (2004). Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs

of illness. *J Occup Environ Med, 46*(6 Suppl), S23-37.

King, N.C.O., Lima, E. P., & Costa, S.E.G. (2014). Produtividade sistêmica: conceitos e aplicações. *Produção, 24*(1), 160-176.

Krol, M., & Brouwer, W. (2014). How to estimate productivity costs in economic evaluations. *Pharmacoeconomics, 32*(4), 335-344.

Krol, M., Brouwer, W., & Rutten, F. (2013). Productivity costs in economic evaluations: past, present, future. *Pharmacoeconomics, 31*(7), 537-549.

Kuhn, G. (2001). Circadian rhythm, shift work, and emergency medicine. *Ann Emerg Med, 37*(1), 88-98.

Lamontagne, A. D., Keegel, T., Louie, A. M., Ostry, A., & Landsbergis, P. A. (2007). A systematic review of the job-stress intervention evaluation literature, 1990-2005. *Int J Occup Environ Health, 13*(3), 268-280.

Lazear, E. P., Shaw, K. L., & Stanton, C. T. (2014). The Value of Bosses. *Centre for Economic Performance, 1318* 1-46.

Lerner, D., Amick, B. C., 3rd, Rogers, W. H., Malspeis, S., Bungay, K., & Cynn, D. (2001). The Work Limitations Questionnaire. *Med Care, 39*(1), 72-85.

Lindegard, A., Larsman, P., Hadzibajramovic, E., & Ahlborg, G., Jr. (2014). The influence of perceived stress and musculoskeletal pain on work performance and work ability in Swedish healthcare workers. *Int Arch Occup Environ Health, 87*(4), 373-379.

Martins, G. A. (2006). Sobre Confiabilidade e Validade. *Revista Brasileira de Gestão e Negócios, 8*(20), 1-12.

Mattke, S., Balakrishnan, A., Bergamo, G., & Newberry, S. J. (2007). A review of methods to measure health-related productivity loss. *Am J Manag Care, 13*(4), 211-217.

Mitchell, R. J., & Bates, P. (2011). Measuring health-related productivity loss. *Popul Health Manag, 14*(2), 93-98.

Munir, F., Houdmont, J., Clemes, S., Wilson, K., Kerr, R., & Addley, K. (2015). Work engagement and its association with occupational sitting time: results from the Stormont study. *BMC Public Health, 15*, 30.

Reilly, M.C., Zbrozek, A.S., & Dukes, E.M. (1993). The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics, 4*, 353-365.

Sadosky, A. B., DiBonaventura, M., Cappelleri, J. C., Ebata, N., & Fujii, K. (2015). The association between lower back pain and health status, work productivity, and health care resource use in Japan. *J Pain Res, 8*, 119-130.

Sahu, S., Sett, M., & Kjellstrom, T. (2013). Heat exposure, cardiovascular stress and work productivity in rice harvesters in India: implications for a climate change future. *Ind Health, 51*(4), 424-431.

Salmond, S. S. (2008). Evaluating the reliability and validity of measurement instruments. *Orthop Nurs, 27*(1), 28-30.

Schultz, A. B., Chen, C. Y., & Edington, D. W. (2009). The cost and impact of health conditions on presenteeism to employers: a review of the literature. *Pharmacoeconomics, 27*(5), 365-378.

Sheehan, K. H., & Sheehan, D. V. (2008). Assessing treatment effects in clinical trials with the discan metric of the Sheehan Disability Scale. *Int Clin Psychopharmacol, 23*, 70-83.

Stang, P., Cady, R., Batenhorst, A., & Hoffman, L. (2001). Workplace productivity. A review of the impact of migraine and its treatment. *Pharmacoeconomics, 19*(3), 231-244.

245

Stewart, W. F., Ricci, J. A., Chee, E., & Morganstein, D. (2003). Lost productive work time costs from health conditions in the United States: results from the American Productivity Audit. *J Occup Environ Med, 45*(12), 1234-1246.

Stewart, W. F., Ricci, J. A., Leotta, C., & Chee, E. (2004). Validation of the work and health interview. *Pharmacoeconomics, 22*(17), 1127-1140.

Ulubeyli, S., Kazaz, A., Er, B. Planning engineers' estimates on labor productivity: Theory and practice. *Procedia - Social And Behavioral Sciences*, 119, 12-19, 2014.

Wahlstrom, J., Hagberg, M., Johnson, P. W., Svensson, J., & Rempel, D. (2002). Influence of time pressure and verbal provocation on physiological and psychological reactions during work with a computer mouse. *Eur J Appl Physiol, 87*(3), 257-263.

APPENDIX 1 – IAPT SCORE:

Questions	Nothing	Little	Regular	Very	Totally
how concentrated and efficient have I felt in the last two hours?	0	1	2	3	4
How tired or sleepy have I felt in the last two hours?	4	3	2	1	0
How productive have I felt in the last two hours?	0	1	2	3	4
How able to make work-related decisions have I felt in the last two hours?	0	1	2	3	4
How confident in my work-related decisions have I been in the last two hours?	0	1	2	3	4
How annoyed or upset during work have I been in the last two hours?	4	3	2	1	0
How difficult has it been to get work done in the last 2 hours?	4	3	2	1	0
How excited to work have I been in the last 2 hours?	0	1	2	3	4
How affected by physical symptoms (pain, dizziness, etc.) have I been in the last 2 hours?	4	3	2	1	0
How satisfied with my work performance have I been in the last 2 hours?	0	1	2	3	4

Final Score: _____

Minimum score: 0 points
Maximum score: 40 points
Productivity Percentage (%) = (Final Score/40) X 100

About the Authors:

Fábio Sprada de Menezes, PhD in production engineering, Federal University of Technology, Brazil.
E-mail: fabio_1711@yahoo.com.br
ORCID
0000-0001-8800-1769
Antonio Augusto de Paula Xavier, PhD in production engineering, Federal University of Technology, Brazil. Email: augustox@utfpr.edu.br
ORCID 0000-0002-3926-3911

Contribution of each author:

Contribution	Fábio Sprada de Menezes	Antonio Augusto de Paula Xavier
1. Definition of research problem	√	√
2. Development of hypotheses or research questions (empirical studies)	√	√
3. Development of theoretical propositions (theoretical Work)	√	√
4. Theoretical foundation/ Literature review	√	
5. Definition of methodological procedures	√	√
6. Data collection	√	
7. Statistical analysis	√	√
8. Analysis and interpretation of data	√	√
9. Critical revision of the manuscript	√	√
10. Manuscript Writing	√	

247

Rev. Bras. Gest. Neg. São Paulo v.20 n.2 apr-jun. 2018 p.232-247