

PYTHEIA: A Scale for Assessing Rehabilitation and Assistive Robotics

Yiannis Koumpouros, Effie Papageorgiou, Alexandra Karavasili, Foteini Koureta

Abstract—The objective of the present study was to develop a scale called PYTHEIA. The PYTHEIA is a self-reported measure for the assessment of rehabilitation and assistive robotics and other assistive technology devices. The development of PYTHEIA faced the absence of a valid instrument that can be used to evaluate the assistive robotic devices both as a whole, as well as any of their individual components or functionalities implemented. According to the results presented, PYTHEIA is a valid and reliable scale able to be applied to different target groups for the subjective evaluation of various assistive technology devices.

Keywords—Rehabilitation, assistive technology, assistive robots, rehabilitation robots, scale, psychometric test, assessment, validation, user satisfaction.

I. INTRODUCTION

ASSISTIVE TECHNOLOGY (AT) devices are being developed every day. AT devices can lessen the impact of disability while increasing independence. As technology innovations appear more often nowadays, the selection of the right assistive technology is considered to be crucial [1]-[4]. The emersion of rehabilitation and assistive robotics and other high-tech and “smart” assistive technologies, along with the numerous features provided, makes even more difficult the matching between the new offerings and the real consumers’ needs. Several crucial issues need to be taken into account during the design phase of a new product in order to be able to penetrate the market. Among others, the cost of the assistive technology, the training needed, the service delivery, the new features, and others are considered critical to the final selection by the consumers [5]-[7]. Despite the promises of the new technologies, the abandonment rate of the used assistive devices is a major concern [8]-[9].

As technology advances rapidly, the problem nowadays has been shifted to the improvement and measurement of the quality of life of each individual even when he/she is facing disability conditions. AT can be effectively used as a mean to improve the independent living of a person facing physical limitations and thus, it is closely related to the quality of life [10], [11]. The right evaluation of needs and functionalities and the measurement of the user satisfaction on the developed

assistive device are crucial to support any valid outcome or to further improve or design new rehabilitation products [12]-[14].

As far as the rehabilitation and assistive robotics is concerned, there is no ready to use, already valid and reliable, instrument for evaluating them either on a pilot phase or a commercial one. After extensive searching in the literature, the most well-known and widely used instruments targeting on evaluating overall user experience with assistive technology are the Assistive Technology Device Predisposition Assessment (ATD PA) [15] and the QUEST 2.0 (“Quebec User Evaluation of Satisfaction with assistive Technology”, version 2.0) [16]. Both instruments try to measure the satisfaction of the end users with achievements in a variety of functional areas, when using assistive technology. The measures have been determined to have good reliability and validity and they have been used in research studies in several populations [17]-[19]. However, they are too generic and there is no evidence, at least to the knowledge of the users, that they can be used to test research or commercial rehabilitation and assistive robotics or even any other assistive device in details (i.e. their individual characteristics and functionalities).

Currently, we are involved in a research EU funded project and we are developing two intelligent active mobility assistance robots for indoor environments that provide user-centered, context-adaptive, and natural support. To this end, the robotic devices integrate several innovative functionalities: (i) they can act proactively by realizing an autonomous and context-specific monitoring of human activities and by subsequently reasoning on meaningful user behavioral patterns, and (ii) they can act adaptively and interactively, by analyzing multi-sensory and physiological signals related to gait and postural stability, and by performing adaptive compliance control for the optimal physical support and active fall prevention. Innovative computer vision techniques with the modalities such as range sensor images, haptic information as well as command-level speech and gesture recognition are implemented. All these modules are incorporated in a behavior-based and context-aware robot control framework.

A major problem faced was the absence of a valid instrument that can be used to evaluate the various characteristics and functionalities of the assistive robotic devices (e.g. autonomous navigation, etc.) in order to meet the real needs of the end users.

The study is trying to address this problem and presents the development and validation results of a new instrument called PYTHEIA that intends to be used for the assessment of any technology device (e.g. rehabilitation and assistive robotics

Yiannis Koumpouros is with the Technological Educational Institute of Athens, Department of Informatics, Ag. Spyridonos, Aigaleo, 12243, Greece (corresponding author to provide phone: +30-6944254112; e-mail: ykoupmp@teiath.gr).

Effie Papageorgiou is with the Technological Educational Institute of Athens, Department of Medical Laboratories, Ag. Spyridonos, Aigaleo, 12243, Greece.

Alexandra Karavasili and Foteini Koureta are with the Diaplasia Rehabilitation Hospital, B’ Parodos Artemidos, Kalamata, Greece.

and devices, etc.), while being able to be adopted and measure users' satisfaction in any individual or special characteristic and functionality implemented.

II. MATERIALS AND METHODS

For the development of the PYTHEIA scale, we followed the steps set out below [20].

- *Conceptualization:* The first step in questionnaire development is conceptualization, which involves defining the subject and the variables to be measured.
- *Questionnaire design:* After the conceptual basis has been specified, a first draft of a questionnaire is to be worked on: appropriate wording, order of questions, and definition of answering categories are now the ultimate tasks. One of the first steps in this phase is to devise a questionnaire structure that respondents will find logical. This step sets down the sequence in which the various topics are presented. The question sequence for each topic is then determined. In the process, the variables defined in the previous phase are translated into specific survey questions.
- *Testing:* Testing a section or the entire questionnaire should not start unless design and wording reach the final version. Basically the questionnaire needs to be tested from three different viewpoints: a) wording of questions/answers, order and structure of the questionnaire; b) problems related to translation, cultural background and harmonization aims; and c) in respect of the data collection mode and the involvement of an interviewer.
- *Revision:* After testing, making revisions to the questionnaire based on the test findings is necessary.
- *Data collection:* When a questionnaire is finished it can be implemented in the field. This is the phase in which data are collected. With the implementation of the survey (either by a pilot study or as the real survey) the iterative process of development and revision is terminated.
- *Process monitoring:* It is important with new questionnaires in particular to keep an eye on what happens during data collection.
- *Evaluation:* Questionnaires are the measurement instruments. Reliability and Validity measures of how well this instruments work.

For the purposes of the study we recruited 30 subjects during the pilot phase and 147 subjects for the final testing of PYTHEIA. The study initiated at the beginning of July 2014 and lasted until January 2015. The subjects were of both sexes, aged from 16 to 93 and users of different assistive devices (e.g. walkers, manual and electric wheelchair, rollator with brakes or not, scooter, orthosis, etc.) for at least one month. All participants should have a score in the Mini Mental State Examination (MMSE) [21] equal to or greater than 17. The Bioethics and Deontology Committee of the Technological Educational Institute of Athens approved the study. All relevant data were recorded (i.e. demographic information, score in the performed examinations MMSE and FIM-Functional Independence Measure, etc.). Moreover, the

participants' individual characteristics (i.e. assistive technology used, frequency of use, pathology, etc.) were also collected. The study took place in a private rehabilitation hospital in Greece and was conducted by experienced physical therapists, occupational therapists, and psychiatrists. All subjects were informed about the study accordingly, and once they agreed, they signed the appropriate informed consent and filled in the questionnaires in a private clinic room in the presence of an experienced occupational therapist. They were administered the instrument in two time intervals (the second time was one week later from the first one). According to [22], the sample size was sufficient for the purpose of the study.

The instrument is divided into two main parts. The first part, comprising 15 items, is related to the evaluation of the assistive technology as a whole, while the second part (items 16 to 20) which can be used as many times is needed in order to evaluate any individual characteristic of the assistive technology (e.g. autonomous navigation, oral commands, etc.). This means that the last five questions have to be used every time a new functionality/characteristic has to be tested (e.g. if we want to explicitly evaluate three different functionalities of an assistive technology we have to use three separate times these questions, one for each individual functionality, etc.). In the first part, the first nine questions (item 1 to 9) were answered by using the 6-point Likert scale: 0-N/A, 1-Not at all satisfied, 2-Slightly satisfied, 3-Moderately satisfied, 4-Very satisfied, 5-Extremely satisfied. Questions from 10 to 15 were answered using the 6-point Likert scale: 0-N/A, 1-Not at all (0% of the time), 2-Sometimes (around 25% of the time), 3-Half the time, neutral (about 50% of the time), 4-Often (around 75% of the time), 5-All the time (100% of the time). Finally, the questions of the second part of the scale (items 16-20) utilize the initially presented Likert scale.

A. Statistical Analysis

The statistical analysis conducted both in the pilot and the final phase as well. For the statistical analysis of the data, we used the Statistical Package for the Social Sciences (IBM SPSS version 19). The assumption of normal distribution of the collected data was tested by using the Kolmogorov-Smirnov test. PYTHEIA was measured against its reliability and validity. For assessing the reliability of the instrument we evaluated its internal consistency, test-retest reliability, and repeatability. Internal consistency evaluates how well different questions (items) that test the latent structure of the instrument should give consistent results. The internal consistency was assessed with Cronbach's alpha coefficient by using the data obtained from the initial assessment. A value of 0.70 was chosen as a threshold in order to check the scale's reliability. The Cronbach's α "if item deleted" was used as an additional evaluation test. The test-retest reliability of the instrument is defined as the degree to which the participants maintained their opinion in the repeated measurements of the questionnaire. Test-retest reliability was evaluated by using the intra-class correlation coefficient (ICC) with 95% confidence interval (CI). The ICC, which is the most suitable

statistical test for the assessment of reliability, ranges from 0 to 1, with 1 indicating perfect reliability. The Cronbach's α and ICC correlations were characterized as follows: 0.00-0.25 = little, if any, correlation; 0.26-0.49 = low; 0.50-0.69 = moderate; 0.70-0.89 = high; and 0.90-1.00 = excellent [23]. Finally, the repeatability is defined as the stability of participants' responses over time, that is, the ability of the instrument to give consistent results whenever it is used. The repeatability is determined by calculating the Pearson's product moment correlation coefficient (Pearson's r) between the initial and re-assessment total scores of questionnaires. The Pearson correlation coefficient values were specified as follows: 0.00-0.19 = very weak correlation; 0.20-0.39 = weak correlation; 0.40-0.69 = moderate correlation; 0.70-0.89 = strong correlation; and 0.90-1.00 = very strong correlation.

PYTHEIA's validity was evaluated by assessing the scale's/subscales' construct validity. The construct validity refers to the degree to which an instrument measures the construct under investigation. The item-total correlations within each PYTHEIA subscale were compared in order to test whether all items of each subscale were related to the same construct. Acceptable construct validity should be indicated by high or excellent (0.70 to 1.00) item inter-correlations for all item pairings. The discriminant validity can be defined as the extent to which the measure being used will give scores and these scores should not be related to the scores attained from an unrelated measure. In the other words, discriminant validity is demonstrated by evidence that measures of constructs that theoretically should not be highly related to each other are not found to be highly correlated to each other. In our study, we performed the discriminant validity test between the different subscales of PYTHEIA.

III. RESULTS

The characteristics of the participants of the final phase are presented in Table I. The mean age of the patients was 62.45 years (SD 19.29), and 55.7% (64/115) were women. The mean FIM and MMSE scores were 104.74 (SD 65.06) and 25.41 (SD 4.01), respectively. An average time to complete the questionnaire was around 7.5 minutes. A range of assistive technologies was evaluated, such as cane, walker, rollator with brakes, scooter, wheelchair, orthosis, and hearing assistive technology.

Variables	Mean	SD
Age	62.45	19.29
FIM Score	104.74	65.06
MMSE Score	25.41	4.01

A. Reliability

According to the analysis conducted, the overall Cronbach's α was 0.793, indicating sufficient consistency (Table II). The various reliability measures are summarized in Table III. The ICC (=0.992) was excellent, indicating that the PYTHEIA total scores were highly consistent between the two occasions (initial assessment and reassessment), whereas the Cronbach α

if item deleted (ranging from 0.747 to 0.808) were also highly consistent between the two occasions and are presented in Table II. The paired-samples t-test between the initial assessment and the reassessment indicated no statistically significant systematic bias. The Pearson correlation coefficient was 0.984, thereby indicating stability of participants' responses over time.

TABLE II
PYTHEIA ITEM ANALYSIS

	Mean (SD)	Cronbach's α if item deleted
Item 1	4.56(0.793)	0.798
Item 2	4.33(1.220)	0.791
Item 3	3.52(2.130)	0.786
Item 4	4.57(1.057)	0.800
Item 5	4.65(0.832)	0.800
Item 6	4.68(0.749)	0.788
Item 7	4.76(0.696)	0.795
Item 8	4.90(0.345)	0.794
Item 9	4.72(0.690)	0.796
Item 10	4.65(0.917)	0.787
Item 11	4.16(1.330)	0.792
Item 12	3.38(1.618)	0.792
Item 13	4.33(1.203)	0.790
Item 14	2.34(2.335)	0.808
Item 15	4.49(1.027)	0.794
IF1	1.58(2.250)	0.748
IF2	1.75(2.367)	0.748
IF3	1.73(2.343)	0.747
IF4	1.72(2.347)	0.748
IF5	1.77(2.382)	0.747

TABLE III
PYTHEIA'S RELIABILITY MEASURES

Characteristics	Measure/test	Value	Significance (p-value)
Internal consistency	Cronbach's α	0.793	
Repeatability	Pearson's r	0.984	0.000
Test-retest reliability at initial assessment	ICC (95% CI)	0.992	0.000
Test-retest reliability at reassessment	Paired-samples t-test		0.059

The aforementioned results verify the initial findings from the pilot study with 30 participants (Cronbach's α =0.770, ICC=0.986, Pearson's r =0.972)

B. Validity

Examination of item convergent validity showed that all item intercorrelations for all item pairings were strong or excellent. Pearson's r ranged from 0.946 to 0.996 for the first dimension "Independent Functionalities", from 0.465 to 0.724 for the second "Fit to Use", and from 0.354 to 0.732 for the third dimension "Ease of Use". This would provide evidence that all dimensions' items are related to the same construct (Table IV).

Examination of discriminant validity is presented in Table V. Since 0.052 and 0.223 is less than 0.85, we can conclude that discriminant validity exists between the "Individual Functionalities", the "Fit to Use", and the "Ease of Use" dimension, respectively. Also, 0.383 is less than 0.85. Thus, we can also conclude that discriminant validity exists between the "Fit to Use" and the "Ease of Use". The three dimensions measure theoretically different constructs.

TABLE IV
CONVERGENT VALIDITY OF PYTHEIA (ITEM-TOTAL SCORE CORRELATIONS)

	Pearson's <i>r</i>
"Individual Functionalities"	
IF1	0.946
IF2	0.991
IF3	0.993
IF4	0.991
IF5	0.996
"Fit to Use"	
Item 1	0.724
Item 2	0.695
Item 6	0.681
Item 7	0.633
Item 8	0.614
Item 9	0.465
Item 10	0.600
Item 13	0.655
Item 15	0.719
"Ease of Use"	
Item 3	0.354
Item 4	0.518
Item 5	0.485
Item 11	0.628
Item 12	0.612
Item 14	0.732

TABLE V
DISCRIMINANT VALIDITY OF PYTHEIA (PEARSON'S R)

	"Individual Functionalities"	"Fit to Use"	"Ease of Use"
"Individual Functionalities"	1	-0.052	-0.223
"Fit to Use"	-0.052	1	0.383
"Ease of Use"	-0.223	0.383	1

The results confirm the initial findings of the pilot phase.

IV. CONCLUSION

The objectives of the study were to evaluate the validity, reliability, and applicability of the PYTHEIA instrument as a scale that can be used to assess rehabilitation and assistive robotics, as well as other technology assistive devices. The target population was composed of heterogeneous groups. The population participated in the study can assure the usage of the instrument in patients with different diseases, ages, and disabilities using various assistive technologies. This can be considered as an asset in our study in order to generalize the outcomes. The next steps will be to evaluate PYTHEIA with patients using two new robotic assistive devices which are being developed in the frames of a EU funded research project.

According to the results presented in the study, PYTHEIA is a valid tool of assessment of satisfaction with assistive technology. The results of statistical analysis support the validity and reliability of the scale because all the items were related to the total score and Cronbach's α and ICC values were excellent, thus indicating that the responses of our sample were internally consistent and stable across time.

As a conclusion, we support that PYTHEIA is applicable and can be proved a valuable tool for the evaluation of rehabilitation and assistive devices. It is essential here to mention that PYTHEIA was tested in the Greek population and all the questions were prepared in Greek. Appendix

presents a draft translation of PYTHEIA in English, without being tested for its validity and reliability. This could be a future study.

According to the results, the developed PYTHEIA can be used to measure the satisfaction of the end users with rehabilitation and assistive technologies, while being able to assess any individual characteristic and functionality of the used device. This can be considered as a valuable and unique asset compared to any other already existing scale.

APPENDIX

TABLE VI
PYTHEIA SCALE

Item #	Question
<i>PART A</i>	
1	Rate your satisfaction with the supporting device and the services provided in relation to the adaptability in the spaces you spend your everyday life (home, work).
2	Rate your satisfaction with the supporting device and the services provided in relation to its contribution to the improvement of your everyday life.
3	Rate your satisfaction with the supporting device and the services provided in relation to the ease of learning all individual functions.
4	Rate your satisfaction with the supporting device and the services provided in relation to the ease of learning the basic functions (the functions that concern me more).
5	Rate your satisfaction with the supporting device and the services provided in relation to the ease of use (complexity, required effort).
6	Rate your satisfaction with the supporting device and the services provided in relation to how secure it is.
7	Rate your satisfaction with the supporting device and the services provided in relation to the dimensions (height, width, length).
8	Rate your satisfaction with the supporting device and the services provided in relation to the weight.
9	Rate your satisfaction with the supporting device and the services provided in relation to if the functionalities existing are sufficient.
10	I will feel more secure (protected, confident) when using this assistive device.
11	I will feel more autonomous when using this assistive device.
12	I will need help from another person to use the assistive device.
13	I will feel comfortable to use the assistive device around the community.
14	I will feel comfortable to use the assistive device among my colleagues (working environment).
15	I will feel comfortable to use the device around friends and family.
<i>PART B</i>	
IF1	Rate your satisfaction with the specific feature of your assistive device in relation the ease of use.
IF2	Rate your satisfaction with the specific feature of your assistive device in relation to the help it provides in your everyday life.
IF3	Rate your satisfaction with the specific feature of your assistive device in relation to how safe/secure it is.
IF4	Rate your satisfaction with the specific feature of your assistive device in relation to its reliability (i.e. whether it applies always correctly).
IF5	Rate your satisfaction with the specific feature of your assistive device in relation to the feeling of safety (I will feel more secure, protected, confident when using it).

ACKNOWLEDGMENT

We thank the physical therapists P. Siavelis and D. Alexopoulou, working in the rehabilitation hospital that helped us in the recruitment process and collection of the data from the participating patients in the study.

REFERENCES

- [1] H. P. Parette, J. J. Hourcade, A. van Biervliet, "Selection of appropriate technology for children with disabilities", *Teaching Exceptional Children*, vol. 25, no. 3, pp. 18-22, 1993.
- [2] Pogoe.org, "Choosing the appropriate assistive device: A card sorting activity | POGoe - Portal of Geriatrics Online Education", 2009. (Online). Available: <http://www.pogoe.org/productid/20465>. (Accessed: 10- Jan- 2016).
- [3] S. Isabelle, S. Bessey, K. Dragas, P. Blease, J. Shepherd and S. Lane, "Assistive Technology for Children with Disabilities", *Occup Ther Health*, vol. 16, no. 4, pp. 29-51, 2003.
- [4] M. Scherer and J. Galvin, "An outcomes perspective of quality pathways to the most appropriate technology", in *Evaluating, Selecting and Using Appropriate Assistive Technology*, M. Scherer and J. Galvin, Ed. Gaithersburg, MD: Aspen, 1996, pp. 1-26.
- [5] S. Winkler, B. Vogel, H. Hoenig, D. Ripley, S. Wu, S. Fitzgerald, W. Mann and D. Reker, "Cost, Utilization, and Policy of Provision of Assistive Technology Devices to Veterans Poststroke by Medicare and VA", *Medical Care*, vol. 48, no. 6, pp. 558-562, 2010.
- [6] G. Dejong, S. Palsbo and P. Beatty, "1. The Organization and Financing of Health Services for Persons with Disabilities", *The Milbank Quarterly*, vol. 80, no. 2, pp. 261-301, 2002.
- [7] WHO. "Joint position paper on the provision of mobility devices in less-resourced settings." (Online). Available: whqlibdoc.who.int/publications/2011/9789241502887_eng.pdf. (Accessed: 12- Jan- 2016).
- [8] I. Ebner. "Abandonment of assistive technology". (online). Available: http://www.florida-esc.org/atcomp/_PDF/MATR%20Abandon%20of%20Assistive%20Technology.pdf. (Accessed: 08- Jan- 2016).
- [9] S. S. Johnston, and J. Evans, "Considering response efficiency as a strategy to prevent assistive technology abandonment", *Journal of Special Education technology*, vol. 20, no. 3, pp. 45-50, 2005.
- [10] M. Scherer, "Outcomes of assistive technology use on quality of life", *Disability and Rehabilitation*, vol. 18, no. 9, pp. 439-448, 1996.
- [11] J. Jutai, "Quality of life impact of assistive technology", *Rehabilitation Engineering*, vol. 14, pp. 2-7, 1999.
- [12] M. Scherer and L. Cushman, "Predicting satisfaction with assistive technology for a sample of adults with new spinal cord injuries", *Psychological Reports*, vol. 87, no. 3, pp. 981-987, 2000.
- [13] G. Craddock, A. Stankovic, T. MacKeogh and M. Scherer, "Online advocacy: user experience and evaluation", in *Challenges for assistive technology: AAATE07*, Eizmendi G, et al., Ed. Netherlands: IOS Press; 2007, pp. 748-752.
- [14] M. Scherer, J. Jutai, M. Fuhrer, L. Demers and F. Deruyter, "A framework for modelling the selection of assistive technology devices (ATDs)", *Disabil Rehabil Assist Technol*, vol. 2, no. 1, pp. 1-8, 2007.
- [15] M. Scherer, *Matching Person and Technology Process and Accompanying Assessment Instruments*. Webster, NY: The Institute for Matching Person & Technology, Inc., 1998.
- [16] L. Demers, R. Weiss-Lambrou and B. Ska, *Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST Version 2.0) An Outcome Measure for Assistive Technology Devices*. Webster, NY: The Institute for Matching Person & Technology.
- [17] M. J. Scherer, Laura A. Cushman, "Measuring subjective quality of life following spinal cord injury: a validation study of the assistive technology device predisposition assessment", *Disability and Rehabilitation*, vol. 23, no. 9, pp. 387-393, 2001.
- [18] M. J. Scherer and G. Craddock, "Matching Person & Technology (MPT) assessment process", *Technology & Disability, Special Issue: The assessment of Assistive Technology Outcomes, Effects and Costs*, vol. 14, no. 3, pp. 125-131, 2002.
- [19] L. Demers, M. Monette, Y. Lapierre, D. Arnold and C. Wolfson, "Reliability, validity, and applicability of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST 2.0) for adults with multiple sclerosis", *Disability and Rehabilitation*, vol. 24, no. 1-3, pp. 21-30, 2002.
- [20] G. Brancato, S. Macchia, M. Murgia, M. Signore, G. Simeoni, K. Blanke, T. Körner, A. Nimmergut, P. Lima, R. Paulino and J. Hoffmeyer-Zlotnik, *Handbook of Testing in the European Statistical System*. European Commission, 2006.
- [21] A. Burns, C. Brayne and M. Folstein, "Key Papers in Geriatric Psychiatry: mini-mental state: a practical method for grading the cognitive state of patients for the clinician. M. Folstein, S. Folstein and P. McHugh, *Journal of Psychiatric Research*, 1975, 12, 189-198.", *Int. J. Geriatr. Psychiatry*, vol. 13, no. 5, pp. 285-294, 1998.
- [22] R. Henson, "Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice", *Educational and Psychological Measurement*, vol. 66, no. 3, pp. 393-416, 2006.
- [23] J. C. Nunnally and I. R. Bernstein, *Psychometric Theory*, 3rd ed., New York: McGraw-Hill, 1994.