Proposing a Perceived Expertise Tool in Business Data Analytics

PANAGIOTIS GERMANAKOS∗, UX S/4HANA, Product Engineering, IEG, SAP SE, Germany
ZACHARIAS LEKKAS, Dept. of Comm. and Media Studies, National & Kapodistrian University of Athens, Greece
CHRISTOS AMYROTOS, InSPIRE Center, School of Sciences, University of Central Lancashire, Cyprus
PANAYIOTIS ANDREOU, InSPIRE Center, School of Sciences, University of Central Lancashire, Cyprus

The business data analytics domain exhibits a particularly diversified and demanding field of interaction for the end-users. It entails complex tasks and actions expressed by multidimensional data visualization and exploration contents that users with different business roles, skills and experiences need to understand and make decisions so to meet their goals. Many times this engagement is proven to be overwhelming for professionals, highlighting the need for adaptive and personalized solutions that would consider their level of expertise towards an enhanced user experience and quality of outcomes. However, measuring adequately the perceived expertise of individuals using standardized means is still an open challenge in the community. As most of the current approaches employ participatory research design practices that are time consuming, costly, difficult to replicate or to produce comparable, unbiased, results for informed interpretations. Hence, this paper proposes a systematic alternative for capturing expertise through a Perceived Expertise Tool (PET) that is devised based on grounded theoretical perspectives and psychometric properties. Preliminary evaluation with 54 professionals in the data analytics domain showed the accepted internal consistency and validity of PET as well as its significant correlation with other affiliated theoretical and domain-specific concepts. Such findings may suggest a good basis for the standardized modeling of users’ perceived expertise that could lead to effective adaptation and personalization.

CCS Concepts: • Human-centered computing; • Information systems; • Applied computing;

Additional Key Words and Phrases: Adaptation, Personalization, Human Factors, User Modelling, Perceived Expertise, Business Analytics, Data Visualizations

ACM Reference Format:

1 INTRODUCTION

Users’ perceived expertise could be considered as one of the most influential individual differences in building effective adaptation and personalization systems and services, but also one of the most difficult factors to assess. Although the importance of its quantification has long been recognized, its inherent structure refers to domain-specific, subjective and qualitative measures that relate to the skills of individuals and their consistent optimum behaviour of actions in terms of efficiency during execution and effectiveness of their results [17]. Hence, perceived expertise may be regarded as a relevant human factor (i.e., relevant-expertise, that its performance does not necessarily need to be assessed in terms of

∗Also affiliated with the InSPIRE Center, Cyprus
an accuracy measure relative to a ground truth [23]) that differs across individuals and in relation to the application 
field they function. Especially, in the business analytics domain end-users, that usually have different business roles, 
backgrounds and statistical knowledge, may engage into demanding data visualization and exploration activities as 
part of their daily responsibilities (e.g., estimate numerical values, predict categorical future outcomes), expecting to 
create an understanding of the information around a task so to make informed decisions and take actions. Inability to 
model adequately their expertise so to provide best-fit data representations or guidance during exploration may lead to 
an overwhelming and frustrating experience for them with poor outcomes. There are many works that studied the 
influence of expertise (or experience) in the adaptation and personalization of data visualizations. Indicatively, [29] 
found that an initial understanding of the data visualization types could be predicted significantly by the business users’ 
experience, in terms of education, working in a data-driven job, and the degree of statistical knowledge. Results which 
have been partially confirmed by [35], arguing that users’ experience, besides preference, also affects satisfaction and 
the capability of being familiarized or switching between graphs to obtain information (e.g. novice users may have 
greater difficulties of using different visualization types). In a similar line of work, [2] supports that users’ skills are 
directly related to their performance when interacting with specific data visualizations and accuracy on respective 
visual tasks in relation to complexity (also supported by [24, 34]).

Nevertheless, these works do not embrace a systematic method to acknowledge and model users’ perceived expertise 
but rather they apply ad-hoc empirical procedures that might be difficult to interpret and reproduce. Given the lack, 
to our knowledge, of a standard tool [15, 22] in the data analytics domain that could facilitate a relevant accurate 
measurement of the end-users’ expertise, this paper proposes a 10-item Perceived Expertise Tool (PET), that consists of 
fundamental psychometric properties of this concept. Underlying concern was to develop a tool that would be driven by 
solid theoretical grounds, as is the five-stage model (i.e., Novice, Advanced Beginner, Competent, Proficient, and Expert) 
of skill acquisition by Dreyfus & Dreyfus [12], indicating how skillful individuals are in the domain they function. 
This model has been extensively applied and adopted in various domains, like medical/ clinical [13, 27], nursing [6], 
engineering [1], or education [26]. The theoretical viewpoints of the model have been applied in situation specific 
scenarios through observation and continuous monitoring, providing the researchers with the contextual information to 
understand and associate subjects to a specific level of skill acquisition. Nonetheless, PET reflects the theoretical mental 
activities and applied mechanics involved during the skill acquisition process, and it is available as an “off-the-shelf” 
self-reporting instrument that may have a wider reach, can be replicated and produce comparable results. Since, it offers 
an alternative approach to time consuming participatory research practices [10], that require long periods of preparation, 
scheduling and execution in-the-field through long-term sessions with focus groups, interviews and observation.

Building a psychometric tool is an iterative process which is based on smaller cycles of components’ design, revision 
and progressive refinement based on the received feedback. Accordingly, in the remaining sections the methodological 
approach and design principles that produced the individual scales of the questionnaire are detailed. Preliminary 
evaluation of PET with 54 data analysts in various business roles, that engage with business data analytics tasks in 
their daily routine, showed that PET has an acceptable internal consistency with most of the items worth retaining, 
and a significant correlation with the affiliated theoretical concepts of self-perceived expertise and self-efficacy as well 
as domain-specific (i.e., data analytics) ones like visual literacy, problem solving and decision making (e.g., Proficient 
and Expert participants have had higher visual literacy and a tendency towards new and more difficult problems and 
decisions). Such findings provide a good basis for a more systematic modeling of users’ perceived expertise, prediction 
of interaction behaviours and interpretations that could lead to effective adaptation and personalization techniques and 
interventions during the data visualization and exploration process based on individuals’ skills.
2 METHOD FOR CONSTRUCTING THE PERCEIVED EXPERTISE TOOL

Creating a tool that would enable the more holistic calculation of end-users’ expertise in the data exploration and visualization domain is a tedious and challenging task. Main concern is to design a construct that will be able not only to inclusively capture through its contents the opinion and behaviour of potential participants but also to explain through a solid theoretical framework their responses and attitudes with respect to the object of investigation, in this case skill acquisition and awareness. Accordingly, following an extensive investigation and literature review, the Dreyfus & Dreyfus five-stage model of skill acquisition [12] was qualified for the purpose of this research and domain of application, for 3 main reasons: (a) It breaks down the process of practical skill acquisition into 5 distinct but inclusive stages making the approach more rigorous and manageable, (b) it is sufficiently associated with the expected soft skills and responsibilities, needs and actions required by the end-users in the business data analytics area for accomplishing data exploration and visualization tasks, and (c) it relates to individual differences in information processing associated with learning, problem solving, and decision making; fundamental mental operations around logic and reasoning that data analysts and related business roles heavily apply in their daily activities. Accordingly, the model is composed of 5 main stages of skill acquisition, namely, Novice, Advanced Beginner, Competent, Proficient, and Expert; and the progression of skill proficiency, i.e., from rigid adherence to rules (Novice) to an intuitive mode of reasoning (Expert), is based on tacit knowledge (i.e., the type of knowledge that is not explicit or formal or it is difficult to transfer to another person by means of writing it down or verbalizing it.) Each one of these stages presents a number of typical characteristics that best describe the skills and usual related actions. In brief, in the Novice stage, a person follows rules or plans as given, without context or exercising any judgement, and with no sense of responsibility beyond following the rules exactly. An Advanced Beginner has limited situational perception, treating all aspects of his work separately with equal importance. At this stage individuals start creating and identifying conditional rules although they do not experience still significant personal responsibility. Competence develops when individuals develop organizing principles, e.g., plans or routines, to quickly access the particular rules that are relevant to the specific task at hand; hence, competence is characterized by active decision making in choosing a course of action while at the same time individuals can cope with multiple activities. Proficiency is shown by individuals who develop intuition to guide their decisions and devise their own rules to formulate plans. Usually, they have a holistic view of a situation, prioritize importance of aspects, and adapt to the situation at hand. Finally, Expert persons have an intuitive grasp of situations based on deep tacit understanding, have a vision of what is possible and they use analytical approaches in new situations or in case of problems [11, 12]. Thus, building upon the aforementioned grounded theoretical perspectives the PET questionnaire has been developed, simulating the five distinctive stages of the model similarly as five progressive levels of skill acquisition. Questionnaires are commonly used in HCI research, providing a convenient way to access users’ perceptions and experiences on specific topics of interest. For the methodological approach the 7 distinct steps for questionnaire construction proposed by Peterson (2000) adjusted for the purpose of this work: [28]:

1. Review literature and requirements (examine academic literature for definitions and dimensions of specific constructs (or global measures of the construct), theoretical models, previously constructed and tested instruments),
2. Analysis of research findings, selection of theory and development of a list of questions or items reflecting the main theoretical principles,
3. Assess prioritized questions or items (screen questions by asking targeted business users to examine potential items against the theoretical definitions of each stage of skill acquisition – verifying the tool’s semantics and controlling the span of different roles and experiences),
Table 1. Perceived Expertise tool – hierarchical items of skill acquisition

<table>
<thead>
<tr>
<th>ID</th>
<th>PET Items (1 = strongly disagree; 5 = strongly agree)</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I strictly adhere to taught rules, theories or principles for analyzing my data</td>
<td>Novice</td>
</tr>
<tr>
<td>2</td>
<td>I trust my judgement for evaluating the current data analysis situation (upon receiving a request and prior starting my task), over the taught practices and models for data analysis</td>
<td>Novice</td>
</tr>
<tr>
<td>3</td>
<td>I often perceive current data exploration situations as &quot;familiar&quot; and I use previous relevant cases to tackle them</td>
<td>Adv. Beginner</td>
</tr>
<tr>
<td>4</td>
<td>As an initial reaction during the preparation phase, I usually treat new data analysis requests as separate cases, instead of classifying them in predetermined categories based on my experience</td>
<td>Adv. Beginner</td>
</tr>
<tr>
<td>5</td>
<td>Before I start a data exploration, I usually have an understanding of the intermediate steps (specific methodology/strategy in mind) that will lead me to my goal</td>
<td>Competent</td>
</tr>
<tr>
<td>6</td>
<td>I have the capability to deal with more than one requests at the same time (multiple activities/requests) during data analysis</td>
<td>Competent</td>
</tr>
<tr>
<td>7</td>
<td>I have a holistic understanding, based on my experience, of the data analysis that is required for a specific request</td>
<td>Proficient</td>
</tr>
<tr>
<td>8</td>
<td>I usually adapt models or data exploration methods to my current data analysis situation, e.g., indicating deviations from normal patterns (that may apply to repeated data analysis requests)</td>
<td>Proficient</td>
</tr>
<tr>
<td>9</td>
<td>I take the risk of creating alternative exploration paths for analyzing data based on my perception of what is possible and more effective</td>
<td>Expert</td>
</tr>
<tr>
<td>10</td>
<td>I trust my intuition during data analysis</td>
<td>Expert</td>
</tr>
</tbody>
</table>

4. Determine types of questions to be asked (open-ended versus closed-ended verbal stimuli; asking people to compare (more or less), endorse (true or false), state the frequency of an action (always or never), comment on the influence of an item, rate the intensity of agreement or disagreement).

5. Decide on specific wording (brevity, relevance, free of ambiguity, i.e., no unfamiliar terms, consider that the modifiers such as “very”, “quite”, etc., and “could”, “should”, and “might” will have multiple interpretations),

6. Determine the structure and size of the questionnaire (visual appearance of the questionnaire; number of items per stage; if using a rating scale, must determine number of categories, labeling and type, like semantic differential (hot... cold), Likert (strongly disagree... strongly agree), or staple (-5... +5)), and

7. Evaluate the questionnaire (for the scope of this work internal consistency and test-retest reliability; validity – internal and external; theory verification – also in relation to complementary theoretical dimensions, i.e., Self-Perceived Expertise, Self-Efficacy, Visual Literacy, Decision Making and Problem Solving (see section 4); and for future work dimensionality – whether the questionnaire reflects a single construct or trait; generalizability – utility of the tool in different research settings and application fields; and theoretical validation/association with other human factors like specific perceptual, cognitive and affective factors).

3 DESIGNING THE PET SCALES

As mentioned earlier, PET is an “off-the-shelf” instrument that examines the perceptual perspective of end-users towards their expertise in the data analytics domain, through self-reporting rather than adhering to more participatory research design principles [10] – consistently involving the research team throughout the assessment process (minimizing the risk for, e.g., biased interpretation of the results by “outsiders”, overhearing end-users or misjudging the importance and weight of their views.) Key characteristics focus on the mobility, applicability and flexibility, complying also with the increasing utilization of remote user research methods in the business sector via digital communication technologies and platforms. Expected benefits lie upon the: (a) Fast adoption and wider reach, e.g., integrate to existing User Research practices in the business domain (e.g., as a post-test survey during usability studies), recruit participants worldwide and distribute it via current user or customer research platforms (e.g., SAP Qualtrics, User Zoom) or to selected mailing lists; (b) profit from the larger volume of collected data (compared to conventional approaches of application, during, e.g.,
Proposing a Perceived Expertise Tool in Business Data Analytics

UMAP ’21 Adjunct, June 21–25, 2021, Utrecht, Netherlands

interviews and observation, focus groups), extracted from a more diverse base of relative business roles and personas (usually organization- and industry-specific) that share the same objectives during the data analysis process; and (c) standardization (ease to compile and analyze), replicability (follow the same process, measure the same variables and control external influential factors during execution, e.g., in focus groups: Dominant personalities, participants are hesitant to voice their opinion freely, whereby at the same time save cost and effort for re-use) and comparable outcome (generate rich data-sets from various user studies increasing the validity and quality of a result).

Consequently, given the theoretical considerations and characteristics of each stage of the Dreyfus & Dreyfus model, a number of items have been designed in a form of a questionnaire that best represent them. Current decision was to comply with the overarching purpose of PET for simplicity, generating currently 2 items per stage (in total 10 items) with the responses to be provided at a 5-point Likert-type scale (see Table 1) and with the total allocation of values to range between 10 and 50 points. Driven by the nature of the theoretical model there was the expectation that skill acquisition would incrementally evolve through the 5 different levels of the proposed instrument, signifying a nested iterative progress across the stages from the Novice to the Expert. This would be achieved by adhering to a continuous theoretical and practical mastery of the qualities expressed by its items, in a way that the previous level would feed the next (progressively disclosing the details (or depth) of the related skill), till there is an accumulated gain of learning and knowledge. Accordingly, a horizontal movement (or coverage) across PET’s hierarchical structure can be observed, obtaining a holistic understanding, knowledge, and progressive skill acquisition while operating on the various goal-directed items (or levels) that present a continuous/gradual growth and development through a specific input (trigger) and a transformed valuable outcome (knowledge). As such, the questions have been intentionally formulated and positioned in a hierarchical order and not randomly, since according to theory individuals should be engaged in a scenario that would be able to employ a more structured and progressive (successive) thinking approach (transformation) while they explore PET’s items (mental functions). With this setting there was also the expectation that the progressive proficiency of skills would be gradually reflected through the cumulative scoring in the 5 individual levels (groups of items) of PET, classified as follows: 10-18 (Novice), 19-26 (Advanced Beginner), 27-34 (Competent), 35-42 (Proficient), and 43-50 (Expert). Finally, it is worth mentioning that the composition and phrasing of the items considered, in a non-exhaustive manner, alleviating any potential biases that questionnaires are keen to create during their application, e.g., respondents avoiding using extreme response categories; agreeing with statements as presented (acquiescence bias or ‘Yes’ bias) [18]; confirmation bias (seeking information that supports pre-existing beliefs or opinions); habituation bias (same pattern of responses to questions that are repetitive or phrased similar); non-response bias [32]; or scoring for seeking acceptance (social desirability bias) [30].

4 PRELIMINARY EVALUATION

4.1 Motivation and Research Hypotheses

The lack of a systematic, standard, instrument to measure perceived expertise in the data analytics sector – and across domains [15, 22], was the main motivation of this work. Tackling an open challenge in user modeling regarding the collection of data that refer to this type of “relative” individual characteristics, and which it mainly resonates into methods that might be proven to be time consuming, costly and not replicable. At the same time may produce outcomes with increased variance — influenced by situation specific factors. Hence, it is very difficult to be comparable for enriching current datasets, that would in turn allow the extraction of more accurate results around perceived expertise,
preferences and behaviors of specific populations on given research objectives. In this respect, this paper proposes PET, investigating its validity through the lenses of the following research hypotheses:

- **H1**: There is a high internal consistency among the various items of PET
- **H2**: There is a correlation of PET with self-perceived expertise and self-efficacy
- **H3**: There is a significant difference between the mean scores of PET’s levels based on the self-perceived expertise, self-efficacy and visual literacy of the participants
- **H4**: There is a correlation of PET with the more rational styles of the decision making and problem solving instruments

### 4.2 Sampling and Procedure

For the preliminary evaluation of PET, data analytics professionals (end-users regularly working with numbers) were recruited using personal invitations. The rationale behind this decision was that the questions under investigation are highly related to their typical real-life tasks involving data and would provide a good indication on the design quality of the tool for the expected target group. Main prerequisite was that end-users have had extensive familiarity – as part of their job profile, with the topic of investigation, i.e., business data analytics, data visualizations and related tasks that require data exploration actions with platforms that produce various types of visualizations. A total of 54 data analysts (31 Female, 23 Male) of average age $M = 33.31$, $SD = 7.53$ participated in the study with average years of experience in the industry $M = 4.82$, $SD = 6.36$. Their actual business roles were ranging from Junior/ Senior Analyst, Retail, Marketing, Advisory Services to Audit, Risk Assessment and Manager. All the users participated voluntarily and provided their consent that their interactions with the Web-system that was hosting the study would be recorded anonymously in the context of an experimental user research study. In addition, the study design followed the internal, external [9], and ecological validity [7] paradigms, allowing the end-users to interact with the study in their preferred time and working environment during the period January 1 - February 15, 2021; refraining of imposing any external influences such as observation or restriction in time availability for answering the questions. To control the learnability effect, participants could undertake only once the study which lasted approximately 20 min.

An electronic environment composed of various Web-forms was created, consisting primarily of the PET questions (see section 3) and other psychometric questionnaires (as control factors) that they are theoretically surrounding the perceived expertise concept. It was expected that by identifying the relationship of their subsequent scales with PET, it would increase the understanding around its behaviour and would be able to interpret the extent of its validity. More specifically, the following additional tools were used:

(a) A self-reporting measure mechanism of perceived expertise, where users could indicate explicitly their level of perceived expertise by simply answering the question: “My level of expertise for the current business role is”, on a 5-point scale where 1 is Novice, 2 is Advanced Beginner, 3 is Competent, 4 is Proficient, and 5 is Expert (reflecting the product of the proposed tool allowing us to further verify the results);

(b) The Subjective Graphical Literacy Scale which is composed from 10 6-point scale questions (1 = not at all good to 6 = extremely good) capturing the perceptions of basic and advanced skills of graph comprehension; experience with tasks involving graphs; and comfort and interest in performing tasks involving graphs [14]. The contents/questions of this tool (e.g., “How good are you at working with line plots?”, or “How good are you at determining the difference between 2 bars in a bar chart?”), seem to be very suitable for the recruited business roles and the business tasks since it aims at collecting information around their expertise in the domain of data exploration and visualization;
(c) A 5-point scale questionnaire (1 = strongly disagree to 5 = strongly agree) for measuring users’ self-efficacy. Specifically, users asked to answer 3 questions “I can adequately handle demanding tasks because I can control stressful and/or unexpected events”, “When I am confronted with a problem, I can usually find a solution if I invest the necessary effort”, and “I can remain calm when facing difficulties because I can rely on my coping abilities.” Self-efficacy is defined as people’s beliefs about their capabilities to produce and perform [4], and it plays a critical role in how individuals think, feel and behave, and for that reason it seems that it is directly connected with the tasks of data analysis, the rational and intuitive aspects of the proposed tool;

(d) The General Decision-Making Style Inventory (DMSI) by Scott and Bruce [31] which includes 25 items and 5 scales (i.e., DM-Spontaneous, DM-Dependent, DM-Rational, DM-Avoidant, DM-Intuitive); and

(e) The Problem Solving Styles Questionnaire (PSSQ) [16] with 20 items and 4 scales (i.e., PS-Sensing, PS-Intuitive, PS-Feeling, PS-Thinking [19]). For last two tools each concept embraces five questions that participants need to respond on a 5-point scale (1 = strongly disagree to 5 = strongly agree). Decision making and problem solving are two processes that frame almost every human action with data visualizations exploration, and analysis. This way would be possible to find some implications that connect the dimensions of PET and reactive responses with behavior mainly used in mental tasks, such as arithmetic operations (e.g. remembering a number in a multiplication problem and adding that number later on) but also an adequate correlation among each other on a theoretical level.

In total, the remote within the sample study was composed of two parts: An introductory part, with information and instructions for the study as well as generic questions around the participants’ age, business role, educational background, years of experience, main tasks, etc., that would help us verify the suitability of the recruited sample with respect to the expected end-user profile and the objectives of this study; and the main part which consisted of the 6 psychometric tools mentioned earlier. The participants could interact sequentially with the environment progressing from one tool to the next answering the questions using the corresponding protocol each time (see above). Given that the validation followed the unmoderated protocol we wanted to ensure that the various items would be understandable by the users who share different experiences and backgrounds, hence we offered optional on-line Q&A sessions during each week where participants could reach out to us for possible clarifications.

4.3 Analysis and Discussion of the Results

At first, the internal consistency was assessed for PET by computing the Cronbach’s alpha. Although, there are no standard guidelines available on expected magnitude for the coefficient, in practice, an alpha greater than 0.60 is considered reasonable in psychological research [20]. Accordingly, the respective analysis revealed the items measuring the various levels of perceived expertise reached a reliability of $\alpha = 0.741$. All items appeared to be worthy of retention, resulting in a decrease in the alpha if deleted. Such a result reveals an acceptable internal consistency of the questionnaire, verifying the first hypothesis ($H_1$). For the sake of completeness, a deeper investigation on the tool’s individual items showed that by dismissing questions 1 and 4 the alpha value could be further increased to $\alpha = 0.748$ and $\alpha = 0.750$ respectively. As such, further iteration on these items will be performed (e.g., rephrasing and increasing the clarity of their contents with respect to the skill acquisition level that they represent), or their possible elimination will be investigated. To formulate a better understanding about the distribution of the sample in relation to PET, we computed the 10 scales of the pertinent items for extracting 5 groups (skill acquisition levels) out of them as discussed in section 3. Descriptive statistics such as frequency distributions and mean indicated that end-users (N = 54) were classified
Table 2. Main correlations between PET, Self-Perceived Expertise (Self-PE), Self-Efficacy (Self-Eff), DMSI and PSSQ styles.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>PET</th>
<th>Self-PE</th>
<th>Self-Eff</th>
<th>DM-Rational</th>
<th>DM-Intuitive</th>
<th>DM-Avoidant</th>
<th>PS-Intuitive</th>
<th>PS-Thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>PET</td>
<td>0.558**</td>
<td>0.399**</td>
<td>0.277*</td>
<td>0.349**</td>
<td>0.281*</td>
<td>-0.286*</td>
<td>0.298*</td>
<td>0.295*</td>
</tr>
<tr>
<td>Self-PE</td>
<td>0.277*</td>
<td>0.399**</td>
<td>0.231*</td>
<td>0.281*</td>
<td>0.281*</td>
<td>-0.286*</td>
<td>0.298*</td>
<td>0.295*</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

...
Proposing a Perceived Expertise Tool in Business Data Analytics

Fig. 1. PET levels mean scores based on Self-Perceived Expertise, Self-Efficacy, Visual Literacy, DMSI (DM) and PSSQ (PS) styles

and Competent (M = 3.55 SD = .680), p < .05, based on their visual literacy; and (c) Expert (M = 4.42 SD = .496) and Competent (M = 3.94 SD = .617), p < .05, based on their self-efficacy (indicating the acceptance of (H3)). These results may imply on one hand that more expert users perceive also themselves as more knowledgeable and skillful in their line of business while on the other hand they may have the ability to understand and work with a wide variety of data visualizations for specific business tasks [29]. Furthermore, stakeholders that present higher levels of expertise and exercise also higher levels of self-efficacy characterized from high self-control and ability to moderate their behaviour to reach their goals. In principle, they have engaged to a bigger number of new and different challenges (in comparison to the other levels), producing successful outcomes and gaining positive experiences. Through this process (i.e., intent, strategies, action, results, evaluation) they have embraced a positive way of thinking, believing that they can achieve the tasks that they set for themselves. Such findings might advise future task-based adaptations with respect to the data visualizations type, complexity and presentation, or data exploration/ navigation support, depending on the expertise level of the end-users [2, 12].

Lastly, as decision making and problem solving are operations that are directly related to the data analytics tasks, we wanted to understand whether there was any association of their individual styles with PET. Therefore, a series of correlations were applied showing statistical significant correlations (see Table 2, and also Figure 1), as follows: (a) positive moderate correlation with DM-Rational (r = .349, n = 54, p = .010), (b) positive weak correlation with DM-Intuitive (r = .281, n = 54, p = .040), (c) negative weak correlation with DM-Avoidant (r = -.286, n = 54, p = .036), (d) positive weak correlation with PS-Intuitive (r = .298, n = 54, p = .029), and (e) positive weak correlation with PS-Thinking (r = .295, n = 54, p = .030) [8]. These outcomes are of particular interest for the validity of PET since it can be observed that it is correlated with the more rational styles of the DMSI and PSSQ instruments (accepting (H4)), verifying the nature of its individual items and applicability in the data analytics area. As it entails more intense rational mental operations for decision making and problem solving during the data analysis process. More specifically, professionals
with dominant on one hand their thinking or rational style try to constantly make decisions through intellectual processes, accepted ideas and values, search for and logical evaluation of alternatives when they make decisions or solve problems. On the other hand, dominance on the intuitive style suggests that they tend to perceive the environment of interaction as a whole, preferring to solve new problems paying attention to detail and a tendency to rely on feeling and experience, while they dislike routine details (a characteristic probably of the Novice or Advanced Beginners levels of PET). A realization that fully complies with the Proficient and Expert levels of PET, where amongst others the element of intuition is intense, and verifies the theoretical framework of Dreyfus & Dreyfus (1980) [12]. In contrast, a strong dominance on the avoidance style dictates a reaction towards avoiding decisions (if possible), a behaviour that again could be hypothesized that fits better to less expert end-users [5, 33]. Related adaptation techniques could regard motivational principles and mechanics for the expert users, or incentives and game-based elements (e.g., rewards, badges) for the less expert ones, so to ensure a consistent engagement on specific data visualization tasks and a positive attitude towards problem solving and decision making in the business data analytics domain.

5 CONCLUSION

This paper proposed a 10-item “off-the-shelf” instrument (i.e., Perceived Expertise Tool (PET)), based on Dreyfus & Dreyfus [12] five-stage model, for measuring the perceived expertise of individuals in the data analytics domain. Accordingly, the methodological approach and design principles that produced the individual scales of the questionnaire have been detailed. Preliminary evaluation with 54 data analysts in the business sector showed that PET has an acceptable internal consistency and a significant correlation with affiliated theoretical and domain-specific concepts (i.e., data analytics) like self-perceived expertise and self-efficacy as well as visual literacy, problem solving and decision making. These findings verify, at some extent, the theoretical validity of the proposed instrument and the ability to predict behaviours that would be beneficial for the creation of related adaptation and personalization techniques and conditions. The afterward, would be the provision of adaptive data visualizations and exploration based on the various levels of perceived expertise of business professionals. Future research, apart from exploring the association of PET with other perceptual and cognitive characteristics, is to address the main limitations of this work referring at re-iterating the items of PET (especially those indicated by the current analysis), repeating the study with a bigger – and more diverse with respect to the 5 levels of skill acquisition – sample for the scope of this research [21], and subsequently perform Exploratory and Confirmatory Factor Analysis for verifying the structure of the relationship among the PET’s variables.

6 ACKNOWLEDGEMENTS

This research is partially funded by the Cyprus Research and Innovation Foundation under the projects IDEALVis (EXCELLENCE/0918/0366) and RABIT (START-UPS/0618/0053) and the European Union under the project SLICES-DS (No.951850).

REFERENCES


[36] Proposing a Perceived Expertise Tool in Business Data Analytics UMAP ’21 Adjunct, June 21–25, 2021, Utrecht, Netherlands