Preprint version: *This is the version that has been accepted for publication and includes author-incorporated changes suggested during the peer review stage. This version does NOT include changed made during the proof-editing last stage before publication of the final version that include minor changes that does not concern the academic content.*
6. July 2018

Spatial decision support systems: exploring differences in pilot-testing with students vs. pilot-testing with professionals

Spatial decision support systems: exploring differences in pilot-testing with students vs. professionals

Romina RODELA [1,2,#], Marta PÉREZ-SOBA[2,3,] Arnold BREGT[2], Peter VERWEIJ[2]

1 Södertörn University, School of Natural Sciences, Technology and Environmental Studies, Alfred Nobels alle´ 7, 141 89 Huddinge, Sweden. Tel: +39 334 9963 658. Email: rominarodela@hotmail.com.
2 Wageningen University & Research, Droevendaalsesteeg 3, 6708 PB Wageningen, the Netherlands.
[3] European Commission Joint Research Centre Ispra, Italy

Abstract

This study explores the implications of engaging students vs. professionals / stakeholders in pilot-testing of SDSS and discusses likely differences in terms of experiences and outcomes for the given pilot-test. To this end we use data collected during two pilot tests of a novel SDSS. The pilot-tests were done with two different groups; one made of thirteen doctoral students, while the other of twelve professionals / stakeholders. The pilot-test served to gather feedback on SDSS usability and other aspects of interest to the development team. Based on the outcomes obtained we develop an analytical framework meant to summarise key aspects impacting on how different (tester) profiles will engage during a pilot, and on feedback they provide. These key aspects include expertise, stage of life, and institutional context (ESI). This framework could offer some help to SDSS / DSS development teams in planning, organizing, and delivering pilot-test, and processing the assessments received.

Key words: Spatial decision support systems, SDSS, DSS, environmental issues, convenience samples, students as research subjects, pilot-testing.

Acknowledgment

This study was prepared as part of Work Package 3 (WP3) of the INVALUABLE research consortium (www.invaluable.fr). The INVALUABLE research consortium is funded under the FP7 ERA-NET scheme, via BiodivERsA, a network of national funding organizations promoting pan-European research that offers innovative opportunities for the conservation and sustainable management of biodiversity (http://www.biodiversa.org).

corresponding author. Please direct queries to rominarodela@hotmail.com

1. Introduction

There has been a substantial growth in the demand and supply of computer-based systems meant to support professionals in spatial planning and environmental decision-making (McIntosh et al., 2011). This trend can be explained by the broadly accepted recognition that environmental issues are complex, and that decision-makers need to balance between multiple, and often competing, claims (Pérez-Soba & Maas, 2015; Rodela et al., 2015; McIntosh et al., 2007). However, research has shown that after spatial decision support systems (SDSS) are made available to the intended end users, these are used little, or not at all (Arnott et al., 2008; Uran & Janssen, 2003). McIntosh et al. (2007) wrote about this and identified a need to expand the research agenda beyond technical aspects and include questions about usability, user needs and tool performance in decision-making processes. Yet, there are challenges to the study of SDSS in the context of real-world decision-making. Anderson et al. (2009) offer an interesting summary of the process leading to their SDSS; they write that it started well with feedback from prospective users, but eventually ended with poor up-take of the SDSS by those who have commissioned it. Anderson et al. (2009) reflect on this and link it to poor communication between those who commissioned the SDSS and the development team, to insufficient user involvement and to insufficient feedback during different stages of that process. They did not have a user involvement and participation strategy, but such a strategy is not common anyhow. In fact, usually the intended users are contacted ad-hoc when the development team needs to have answers to questions, information, data, or needs to pilot-test the SDSS (Poch et al., 2004).

On the other hand, user availability is also an issue. End users of SDSS most commonly are practitioners and decision-makers with busy agendas and, despite the need for a SDSS, might not have the time or the capacity to join at meetings, or sit in pilot-test sessions. As a consequence, when pilot-testing is needed, developers and researchers often reach out to students since these are relatively easy to recruit. The involvement of students for pilot-testing of decision support systems has been reported in a number of publications to include Arciniegas et al. (2013), Ben-Zvi (2012), Giupponi (2007), Gorsevski et al. (2013), Vahidov et al. (2014), and Van der Wal et al. (2016). However, there are methodological issues that come with the involvement of students which have not been discussed by current literature. For instance students might be asked to participate to these activities as part to course requirements, might be asked to role-play, or to simulate an imaginary scenario, about which they have very little, or no, knowledge. Yet, it is not discussed, nor understood, how such circumstances might influence their engagement with the SDSS, and the feedback provided. To the best of our knowledge this topic has not received much attention.

It is important to mention that the recruitment of students as research subjects to test theories, methods or tools is common across a range of disciplines as is management studies, psychology, etc. (Hoodge, 2010; Peterson & Merunka, 2014; Peterson 2001). In psychology, where human behaviour is being investigated, students represent an accessible and convenient on-campus option to which the researcher has easy access. Such samples are referred to as **convenience samples.** However, since students have specific socio-economic and psychological characteristics some scholars are sceptical about the extent to which students can meaningfully participate in research that aims to generalize to a larger population. Recent failures in replicating results of key psychological studies, known as the replication crisis, sparkled a vivid discussion about the use of student samples (see: Bohannon, 2015; Bower, 2016; Van Bavel et al., 2016). That debate offers insight on the implications that arise from recruiting students also when pilot-testing SDSS that we consider later in the analysis. We also acknowledge a few studies in adjacent fields to SDSS where differences between students and professionals in the type of feedback provided is

referred to. For instance, Van der Wal et al. (2016) observed differences in how the two groups perceived and interacted with their model. Van der Wal et al. (2016) write that students were more likely to converge towards agreement compared to professionals, but also that students discussed the model feedback in relation to the game played (i.e. virtual role playing), while professionals discussed the model feedback in relation to margins, risk and how the (actual) river is likely to behave in the future (i.e. real circumstances).

It is the objective of the present study to explore and raise questions about the implications that arise from pilot-testing SDSS (meant for professional use) with students. To this end we use data collected during two pilot tests of a novel SDSS. The pilot-test workshops were done with two different groups; one made of 13 doctoral students, while the other of 12 professionals / stakeholders. In the next section we first provide information about research methods used and about the two pilot-test workshops. Then, we present and discuss the results obtained and in building upon these we propose a simple analytical framework. Chapter Five closes the study with recommendations for research and practice.

3. Methods

This is an explorative study where we engaged in an iterative process. Thus, we started from a very practical need to pilot-test a novel SDSS, commissioned to members of our project team, and were soon confronted with the challenge of recruiting "testers" with profiles similar to those of prospective users (i.e. professionals, decision makers). Due to recruitment difficulties we then engaged students in a first pilot-test workshop run in 2013. However, aware of the professional needs this SDSS is meant to serve, questions were raised about how and on what aspects students can offer feedback useful to develop a SDSS which is able to meet professional needs. Our team then agreed on the need to recruit professionals for a further pilot-test. In October 2014 the SDSS was tested again with a group of 12 stakeholders who were all professionals.

Student sampling is a topic that in the past years has received attention, mostly in relation to the replication-crisis in psychology. We were broadly aware of the issues associated with this and the first author surveyed relevant literature then used to identify areas of general scientific concern (i.e. statistical reliability, generalizability) as well as narrower potential concerns for the SDSS applicative domain (i.e. implications of role-playing, testers lack of expertise). That led us to formulate assumptions about when, and how, a pilot-test with students would be suitable, given the purpose SDSS are meant to pursue.

Here we like to clarify that the purpose of the pilot-test workshop we did, and report upon, was to pilot test the SDSS and not to test hypothesis about student samples. The latter emerged as a topic of concern at a given step of the SDSS development, which after reviewing literature on student samples, and upon the analysis of data collected (questionnaire and observational notes), led to the formulation of a framework meant to summarise assumptions, and capture core aspects, that we believe have a role in how different groups /sample types would engage in a SDSS pilot-test.

3.1 Background information about the (pilot-test) workshops

QUICKScan is a participatory spatial modelling method designed for group use that links user´s knowledge and preferences to available spatial- and spatial-statistical data (details: *citation anonymized* et al., 2016). The QUICKScan was commissioned by the European Environmental Agency to members of our project team with the intent to support decision-makers in complex decision-making processes where spatial and other type data need to be taken into account.

The purpose of the two workshops was to pilot test the SDSS and gather feedback on SDSS usability, but also since the SDSS is meant to be used in a group context, there was a need to assess if, and how, the SDSS supports learning and helps the participants to share and integrate different knowledge. The QUICKScan workshops are informed by participatory methodology and input from all participants is sought; these are asked to work in small groups and develop model input matrices then used to build scenarios and define causal relationships between drivers and impacts (*citation anonymized* et al., 2016). During the small group activities as well as whole group discussions participants share views, knowledge, discus the indicators and relationships between variables chosen.

A first workshop took place in April 2013. This was done with thirteen doctoral students who took part in an Impact Assessment International Summer School in Scotland (UK). The topic was "green infrastructure", the geographic scope were EU Member States and the session was done as "a role-play" with students simulating to be European public officials. The second workshop took place in October 2014 with twelve professionals who have stakes in the issue discussed, and acted as representatives of stakeholder groups. These professionals - whose profiles included farmers, entrepreneurs, forest estate managers, park managers - are members to a partnership working on initiatives for the regeneration of a rural area in Northern Scotland (UK). The topic was "land use scenarios" and the geographic scope was a local area. The QUICKScan team had no influence on the selection of workshop participants'; in the case of the first workshop participants were selected by the Summer School organizers, while for the second workshop a local gatekeeper invited all members of a local partnership and eventually twelve showed up on that day. Details about the two groups of participants are summarised in Table 1.

These two QUICKScan workshops are comparable on delivery (facilitated session with a moderator and a QUICKScan expert operating the tool), objective (collaboratively producing and discussing different scenarios), length (one day made of a full morning and afternoon) and purpose (demonstration with *no* formal obligations for decision-making), but differ in terms of participants profile (professionals *vs*. students), and stakes participants had in the topic being discussed (real stakes *vs.* playing a role). Both workshops foresaw that participants worked on alternative scenarios, however, the way this was delivered differed due to the participants' profiles. It was expected that students would not be very familiar with the subject matter and for this reason a role-play simulating European decision-making was prepared by development team. On the other hand, professionals were expected to be familiar with the subject matter and were asked to contribute with own expertise and opinion. All were also stakeholders who were often involved in decision-making processes about local affairs. Given that QUICKScan was developed to support professionals in participatory environmental decision-making processes, it is useful to note that the group made of professionals approximates better the profile of the QUICKScan prospective users.

3.2 Data collection and analysis

A questionnaire was developed and used at both workshops (*citation anonymized*). The questionnaire is made of two parts; one is to be administrated before the workshop and a second one is to be administrated after the workshop. Each was filled out by participants themselves independently, the first part in the morning before the session started, while the second part was filled after the session ended.

This data was triangulated with further qualitative data inclusive of; own notes about group dynamics and general comments given by the participants, questions participants asked about the SDSS and visuals, and other peer-to-peer verbal exchange. These notes were summarised and used to interpret and cross-check questionnaire data.

4. Results

A total of twenty-five participants have answered questions on our two-part questionnaire. Participants to the first pilot-test are referred to as group 1, while participants to the second pilot-test are referred to as group 2 (Table 1). Participants to group 1 were students enrolled at PhD programs and reported not knowing each other from before, while to group 2 are professionals and some knew each other. On the basis of secondary data we also know that organizations represented by the participants of group 2 are collaborating in a partnership focused on the regeneration of that rural area.

The data summarised in Table 1 and Table 2 indicate differences between the group made of students with that of professionals / stakeholders in terms of prior knowledge of GIS, completed education, working experience and other aspects of interest. With the pre-test questionnaire we gather that 9 out of 13 students from group 1 have participated to workshops before and are well familiar with GIS, while only 4 out of 12 professionals from group 2 stated prior familiarity with GIS (Table 1). On the other hand data collected with the post-test questionnaire (Table 2) shows that the students from group 1, scored higher on several of the questions targeting knowledge acquisition compared to the professionals from group 2. For instance, students scored higher on understanding the maps produced with the QUICKScan and also scored slightly higher when asked if maps were useful, and when asked if the process to produce the maps was comprehensible to them (Table 2). The variation in the scores is not large, but it points at differences which we assume arise from familiarity with GIS and thus easiness in understanding the maps, tables, and graphs produced.

Table 3 and Table 4 summarise descriptive feedback participants have written about the tool, the workshop, the visuals, and reflect on the learning processes they experienced.

	No. of participants	Age (Mean value)	Education completed	Years of working experience (Mean value)	Familiarity with GIS (Yes, No)	Participate to similar events (Yes, No)	Know other participants from before (Mean value)	Affiliated with.*	Past collaboration*	Does she/he have a direct stake in the problem issue discussed*
Group 1: PhD students	13	30,3	13 MSc	4,2	all	9 out of 13	0	PhD candidates at universities	No	None
Group 2: professionals	12	46	2 none in particular 6 with BSc 3 MSc 1 PhD	27,4	4 out of 13	10 out of 12	7	ENGO; park management authority; association; local trust; public institutes'; farmers alliance; entrepreneurs;	Yes	Yes all 12

Table 1: Summary of sample characteristics

Note: Information marked with * is extracted from reports. All other data is self-reported with the pre-test questionnaire.

	Have you shared own knowledge	Have you acquired new knowledge	Have you up-graded knowledge already have	Has the tool helped to verbalize knowledge	Has the tool helped to share knowledge	How much of own knowledge is integrated in the maps produced	Have you understood the outcomes/ maps produced	Were maps useful	Was the process to produce the maps comprehensible to you	Are you satisfied with the outcomes	Has your position changed
Group 1: PhD students	Mean =5,6 SD = 0,768	Mean=5,79 SD=0,751	Mean=5,79 SD=0,947	Mean=5,4 SD=0,870	Mean=5,5 SD=0,877	Mean=5,6 SD=0,650	Mean=6,18 SD=0,760	Mean= 6,3 SD=0,651	Mean= 6 SD=0,603	Mean=6,2 SD=0,718	Mean=4 SD=1,683
Group 2: stakeholders	Mean=4,9 SD=,515	Mean=5,4 SD=1,04	Mean=4,9 SD=1,32	Mean=4,6 SD=1,311	Mean=4,6 SD=1,443	Mean=4,2 SD=1,267	Mean=4,85 SD=1,357	Mean=5,85 SD=1,357	Mean=5,3 SD=1,215	Mean=4,8 SD=1,215	Mean=3,6 SD=1,084

Table 2: Data for selected questions on the post-test questionnaire

	Do you have any comments or suggestions on how the QUICKScan tool was used?	Do you have any comments or suggestions on how the workshop was organised?	Do you have any comments or suggestions on the type of information/data/ visual tools that have been provided?
Group 1: PhD students	• Maybe [have access to details] a case in advance	• Pity we didn't have more time. • Training & time. • It was not enough time. • In my opinion, it is not very clear, either do diagnose or scenarios.	• Fun! • Maybe some screenshots on a bigger screen.
Group 2: stakeholders	• Larger scale scenarios give better impressions of proposed changes. • Examples showed how it may be used but don't know if time involved in using software will be worthwhile. • It is complicated. • Good to use examples to break out session. • Great potential. I am impressed by it, but concerned about its practical application. • It was just an example but I can see how it could be used more	• Well organized, well presented. • Well organized. Not sure of what value was (had) from after lunch session. • Good. Well-structured day that kept to time. • Morning break would have been good. • Good balance between the sessions. • Well organized, well thought through. • Ok.	• Estate wide scenarios gives good overall impressions, but more detail of individual areas would real decisions to be made • No, it was all fine once I sort of got my head around the spider diagram. • Tool has great potential but only when used well. • Bit overwhelming. • P. was very slick with the QUICKScan tool, shows how useful it can be in true right hands. • Ok.

Table 3: Feedback on tool, workshop and visuals

Note: All answers obtained are reported here. Not all respondents have provided answers to all open questions.

	Please describe in your own words what contributed positively towards your learning experience and in which ways this occurred.	Please describe in your own words what contributed negatively towards your learning experience and in which ways this occurred.
Group 1: PhD students	• Discussion, visualization. • Tool is excellent, maybe need to improve my skills. • Interaction with other participants and discussion/participatory discussion system. • Talk with the others. • Group work, taking over a politician's role. • Visualization, participation and discussion. • Visualization is strong point of QS as well as spatial explicacy. • Facilitation of aims setting in relation to our role and support of the discussion.	• Too little time. • Confused to give score. • Nothing. • Lack of time - not enough to discuss all the options. • Too less time for the group work, especially in task 1.
Group 2: stakeholders	• The ability to see conflict areas helps to decide on future decisions and change. • Demonstration of software and using examples. • Good stakeholder engagement. • The hands on demonstration of Quicksan. • Use of Quickscan as a means of showing openess and clarity of thinking on how a decision may have been reached. • Demonstration of what Quickscan can do. • How it was put across. • This is a good method of identifying and defining issues and potential conflicts between various interests. • Seeing how QuickScan works; apparent case of use, seeing scenarios, that it can handle a range of data. • The iterative process of discussing various benefits (profile) with others. • Great to see people talking dispassionately.	• Seems quite a complicated procedure for producing maps, training be required. • Slightly confused delivery. • After lunch focus group was vague. • The system is complex and would require a dedicated person to input and help with facilitation in getting rules correctly answered. • Discussion of concepts model confusing - doing a worked example would be more beneficial. • Very complex computer system. • Hands on would be useful. • Overcomplicated - to many variables need to focus on conflict areas or benefits. • None.

Table 4: Feedback on learning.

Note: All answers obtained are reported her. Not all respondents have provided answers to all open questions.

5 Discussion

In this section we compare and discuss the data collected in the attempt to consider if, and what, are core differences in the type of engagement with the SDSS, topics discussed and feedback provided across the two groups of pilot-testers.

Results collected with the questionnaire show that the two groups differ in prior knowledge of GIS and years of completed education. These differences are not understood to be an issue *per se,* SDSS are developed for a range of user profiles from policy-makers to practitioners who may, or not, be familiar with GIS. However, of an interest to this analysis is to consider if, and how, familiarity with GIS influences the way pilot-testers engage with the SDSS and the outcomes produced. In our case, feedback provided on the QUICKScan and on the visuals (Table 3) suggests that a circumstance might have occurred where students by virtue of their familiarity with GIS coupled with the fact of being engaged on a general scenario where they had no stakes was perceived as more straightforward experience. Students were engaged in a role-play and well aware of what is expected from them. During that workshop critique of parameters or scenarios was absent and students were busy implementing the instructions the development team handed out. They appeared **task-oriented.** This, we assume, might have contributed to higher level of satisfaction with the maps, tables, and graphs they knew needed to be produced. Contrary to this professionals were asked to contribute with information and knowledge they have about environmental issues of places they know well and care about. They sat in the pilot as experts and representatives of local stakeholder groups. They had comments and questions. For instance a few professionals, with no familiarity with GIS, struggled to understand some of the outputs and questioned the parameters the QUICKScan team has chosen to develop the scenarios. After a question was raised often a few others followed adding further information and knowledge about that matter, or related local circumstances. This contributed to advance the debate on a given issue. For some of the professionals/ stakeholders the outputs produced were not seen as a useful, or an accurate, representation of local circumstances and that critical view was noted also in the feedback provided on maps, tables, and graphs produced. Differences in the perception of the pilot test are noted in the answers summarised in Table 3 and Table 4. In overall, the pilot-test with professionals involved a more critical and lively discussion and professionals appeared to be **issue-oriented,** compared to students who conversely appeared to be more **task-oriented**.

We suggest here that differences we observed across the two groups in terms of engagement, questions raised, and the feedback provided, are linked to their profiles and the implications that arise from that. For instance the implications that arise from participants' expectations towards an SDSS for how they experience and evaluate it. Russo et al. (2018) have reported on the role of expectations and liked that to issues with DSS up-take by practitioners. Here, while considering literature on student samples and SDSS, we elaborate on those differences we observed during the two pilot tests.

Literature on student samples emphasizes that when these are young adults undergoing education most often are living a stage of life when they might not have very well crystalized attitudes, sense of self, opinions and for this reason are not suitable for recruitment in given research projects (see: Hoodge, 2010; Peterson, 2011). We trace evidence supporting these assumptions in our research since we recorded evidence of the two groups engaging and behaving differently during the pilot-test. For instance students diligently followed instructions the development team handed out. On the other hand, professionals had questions about content, expressed disagreement on a few aspects when things on screen were not as they expected or were hard to follow. Professionals are at a more advanced stage of life where they have matured professional and personal experience, and have seen how a professional

environment works. They felt confident and entitled to be critical. Differences in how the two groups experienced the task given to them are seen also in the answers recorded in the post-test questionnaire. The answers students gave suggest the pilot-test was an enjoyable experience and an opportunity to learn where they acted as those on the *receiving* side vis-à-vis those on the *giving* side i.e. the researchers / the QUICKScan team.

While stage of life is assumed to explain why professionals are more likely to voice critical views compared to students there are further circumstances that contribute to influence how participants with different profiles will engage during a pilot-test. We assume expertise also has a role. Professionals in our pilot-test reported several years of working experiences, to be well familiar with the geographical area under discussion, the environmental and social issues present there. They brought into the discussion a fine level of detail, used real examples from that natural area when commenting on aspects of their interests and this in a way that their input was understandable to others in the room. For instance, during the discussion of a fishing scenario (where land use data and accessibility were mapped) the representative of a local water alliance challenged variables appearing on the screen as not represent well water bodies in that area. In that comment difficulties with quantification when rivers appear as lines we expressed: "*most of the data is land based and so it does not speak very easy to me*". That observations sparked a vivid discussion about usefulness of the chosen variables for the intended scenario, and affected the process in given ways. We assume that this level of engagement, with the variables and scenarios, was possible mostly because each participant was also an expert and so able to consider how these perform in practice. The fact that participants were all so intimately familiar with the geographical and socio-economic conditions of that natural area had value and allowed them to reflect how those abstract ideas would perform in the geographical area we worked with. The expertise participants had informed their comments, the type of questions they asked and help them to assess the scenarios and other visuals produced. A circumstance alredy Van der Wal et al. (2016) observed in their workshops as well.

In addition to the above we also assume that (institutional) context will contribute to how participants with different profiles will engage during a pilot-test. In both workshops the team explained that the session is meant as a demonstration of the SDSS and there will be no implications for decision-making. Yet, professionals/ stakeholder were quite concerned with the accuracy of maps and graphs produced. As seen from Table 2 standard deviation the stakeholder sample is high which we have linked to the low scores given by those who were not satisfied with the maps and graphs. We reached out to three participants (outliers) with a telephone interview and were told that to their opinion, maps were not accurate, and the session failed to integrate all relevant variables and capture important things (e.g. mapping of water bodies). It is thus interesting to observe that even if it was understood that this was a "demonstration" session with *no formal* obligations, professionals felt to be there in representation of given areas of interests (tourism, water management, nature protection, etc.) and acted on these throughout the session, raising questions, expressing concerns and disagreement. While they might not had expectations towards what a SDSS should look like, their professional mandate influenced the way they engaged in the discussion and in the assessment of the outcomes produced. They expected accuracy and professionality. On the other hand, students simulated to be policy-makers, did not knew each other from before and performed roles on the basis of instructions received. The fact they were not able to feedback on how the SDSS would accommodate for needs and requirements of the simulated institutional context is not, however, a minor issue as prior research has shown this impacts on SDSS up-take (Anderson et al., 2009).

To conclude, the data collected suggests that there are certain implications that arise from differences in the participants profiles' which seem to have certain implications on how they engage, experience, and

assess the pilot-test. In our case students appeared more task-oriented and satisfied with the outcomes produced, compared to the professionals who appeared more issue-oriented and less satisfied with the outcomes produced. On the basis of the results obtained in the next section we would like to suggest an analytical framework meant to capture key aspects distinguishing students (i.e. young adults) from professionals which might likely influence the performance and assessment of these profiles of a SDSS pilot-test.

6. An analytical framework summarising differences between students and professionals that matter for SDSS pilot-testing

Based on the research reported above, and by taking stock of academic literature on SDSS as well as literature on student samples, in this section we now advance a framework which suggests that participants' profiles matter and will likely have an influence on the pilot-test (this for SDSS meant for use by professionals). This framework considers differences between two profiles (i.e. students and professionals/stakeholders) in terms of **expertise, stage of life,** and **institutional context** and conceptualises how these impact on participants' engagement, personal experience and assessment of the SDSS (Figure 1).

The first aspect of interest is **expertise**, that is understood to be skill in a specific domain (Ericsson et al., 1993, 2007), developed in result to years of work in a given subject area. Long term work and engagement with a given environmental issue is important to develop an in-depth understating of the problem domain, but also facilitates understanding how abstract ideas and theory perform in practice when applied to address that environmental issue. Research has shown that on job training which includes repeated exposure and trial-and-error is key to mature expertise (Ericsson et al. 1993, 2007). Students still undergoing academic training do not have this opportunity. They are asked to understand and/or memorize notions, ideas and theory, and read case studies. They do not have the opportunity to actually work on and apply theory to environmental issues in practice. We acknowledge that while students do well when asked to discuss material in abstract terms, they might lack an in-depth understanding of how ideas and theory can be used to address real-world environmental issues and of the challenges attached to this. On the other hand, professionals have expertise which they can apply during a session and then give feedback on how well a SDSS, or selected aspects, accommodate for given professional needs and how well the outcomes produced match their (professional) expectations.

A second aspect of interest is **stage of life** as students undergoing education might be more likely to perceive this as an opportunity to learn and consequently undertake the task with less critical insight and confidence to voice critical views compared to professionals. Also, as students are undergoing training they are used to demanding cognitive exercises that also is circumstance likely to influence how they perceive, and react to, SDSS functionalities and outcomes they not understand that well. Thus, students might see the need to put extra effort to follow and grasp what is being done during the session as normal and might not feel compelled to report on that in detail after the session ends (e.g. in the questionnaire, or report). Mature students, and particularly those at a **stage of life** when they have some years of work experience and have matured expertise, could be an exception to this, but on the hand they would lack links with the working environment where the SDSS is meant to be used.

A third aspect of interest is the (**institutional**) **context** to which these two different groups belong. We assume that professionals who have a position within their institution, and therefore have a mandate and responsibilities when interfacing with others about a given topic, while students would not. Students

undergoing education are generally be unaware of the needs, issues and pressures professionals face in a professional context on a daily basis. This is a further reason why students are often involved in pilot-tests on the basis of a role-play script prepared by the development team. Also, students might participate as part to a training, programme, or course requirements, and the session most often would involve a (simulated) circumstance they might have learned about in class, or was described in a text. All these circumstances contribute to shape certain expectations and consequently influence the way students engage during the session and then assess the SDSS. Research has shown that SDSS up-take is linked to how well a given SDSS is able to fit into on-going institutional procedures (McIntosh et al., 2011; Merrifield et al., 2013), but students would not be able to comment on that. Students' choices and engagement is not influenced by mandates and stakes in need to be defended, or needs to be voiced in relation to what a SDSS is meant to do. Contrary to this professionals are faced on a daily basis with complex tasks and are aware that a key task of SDSS is to help with demanding decision making. Professionals have background insight needed to comment on how well the SDSS performs on this, and how well it would fit in an organizational context and corresponding institutional practices.

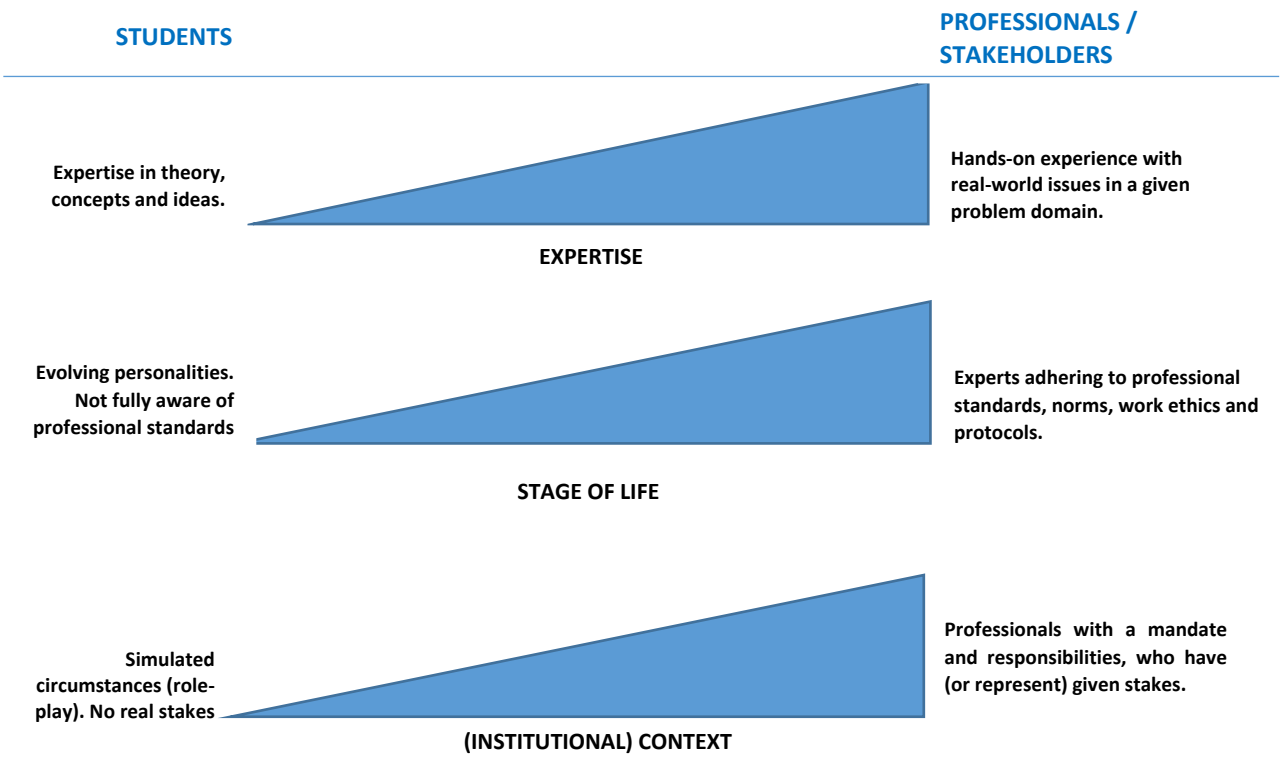

Figure 1: Expertise, Stage of life, and Institutional context (ESI): a simple analytical framework

While the framework seeks to conceptualise differences across participants' profiles in terms of **expertise, stage of life,** and (**institutional**) **context**, it is not its purpose to advocate for the abandonment of student samples in all cases. There are a series of different steps SDSS development teams need to undertake and facilitate starting from an initial conception, to design, development, test, and re-test. Each phase comes with given needs in terms of data, information and feedback the development team needs in order to advance with success towards a final SDSS and its up-take by the intended users. In pointing to differences

between the two profiles this framework can support development teams in thinking over the type of feedback they need during different stages and then, on the basis of this, make appropriate choices. For instance, while professional / stakeholders are needed early on to help the team in identifying the problem domain, to inventory their needs and knowledge, and to set objectives, students with a comparable level of mastery of GIS, and related, could be engaged in preliminary assessments on selected aspects on which the team needs a fast feedback. Also, students could be involved in testing the plan for the pilot-test delivery. They can offer feedback on the pilot process delivery and on the tools designed to assess it (e.g. questionnaire, observation protocol). Also, when it comes to visuals, maps and related students could offer useful feedback. Further, in the case of SDSS meant for group use where moderation is needed, students could be of help in testing the "plan for participation and engagement" meant to allow meaningful involved of all participants.

7. Conclusions and future research

SDSS meant for group use have the objective to support group activities (i.e. exploration, discussion, decision-making) and to this end seek to link knowledge, and preferences, of those in the attendance of the activity to available spatial- and spatial-statistical data. Research has shown that in such cases decision-making will be likely affected by inter-personal interaction which includes (verbal and non-verbal) communication and the social influence exerted by statements, opinions and facts shared by the participants as they interact with the SDSS and with one another (Sambamurthy & Chin, 1994; Zapata, 2009). Yet, when there is a need to pilot-test a newly developed SDSS this is often done with students which limits the possibility to understand how well the SDSS fits expectations, needs and institutional requirements of professionals. In this study we have discussed issues around the use of student samples, and in comparing data collected from two pilot-tests, have shown this might have implications for the type of engagement and feedback students are likely to give.

Based on our exploratory study we suggest a framework where *expertise*, *stage of life* and *institutional context* (ESI) are conceptualised as those factors understood to have an important impact on how testers are likely to perform at, and assess, a SDSS meant for professional use.

Our results show students being more task-oriented and satisfied with the outcomes produced, while the professionals more issue-oriented and less satisfied with the outcomes produced. Professionals/ stakeholders placed the SDSS, and the outcomes produced, in a professional context linking it to needs and issues they face at work on a daily base. As a consequence they were more wary and critical with regards to the SDSS, and its functionalities, compared to the students. Also, results show that professionals were better able to discuss the environmental issue at hand given their understanding of local circumstances and environmental conditions. As mentioned ours is an exploratory study and our conclusions take into account the limited sample size we had. Future research might try to test these observations further and contribute to the debate on student involvement in SDSS pilot-testing.

To conclude, while samples made of students are often seen as a convenient alternative when potential users are difficult to engage, students should not completely replace the intended end users. Instead, student samples could be considered for testing selected aspects of SDSS as for instance a general technical features, which however should be further developed, tested and re-tested in close collaboration with the intended users, or alternatively with comparable professional profiles. This might help to create a sense of relevance, ownership over the final outcome, and facilitate the integration of specific needs, and demands that users have. Our study suggest that pilot testing SDSS with professionals/ stakeholders is essential for a realistic assessment of a SDSS.

8 References

Anderson, C., Beazley, K., & Boxall, J. (2009). Lessons for PPGIS from the application of a decision-support tool in the Nova Forest Alliance of Nova Scotia, Canada. *Journal of Environmental Management* 90, 2081-2089.

Arciniegas, G., Janssen, R., & Rietveld, P. (2013). Effectiveness of collaborative map-based decision support tools: Results of an experiment. *Environmental Modelling & Software,* 39, 159-175.

Arnott, D., Pervan, G. (2008). Eight key issues for the decision support systems discipline. *Decision Support Systems*, 44, 657-672.

Ben-Zvi, T. (2012). Measuring the perceived effectiveness of decision support systems and their impact on performance. *Decision Support Systems,* 54, 248-256.

Bohannon, J. (2015). Reproducibility. Many psychology papers fail replication test. *Science* 349(6251): 910–911.

Bower, B. (2016). Psychology's replication crisis sparks new debate. *Science News*. 189 (7):8.

Ericsson, K.A., Krampe, R.T., & Tesch-Römer, C., (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100, 363-406.

Ericsson, K.A., Prietula, M.J., & Cokely, E.T., (2007). The Making of an Expert. *Harvard Business Review,* 7/8, 115-121.

Giupponi, C., (2007) Decision Support Systems for implementing the European Water Framework Directive: The MULINO approach, *Environmental Modelling & Software*, 22(2): 248-258

Gorsevski, P.V., Cathcart, S.C., Mirzaei, G., Jamali, M.M., Ye, X., & Gomezdelcampo, E., (2013). A group-based spatial decision support system for wind farm site selection in Northwest Ohio. *Energy Policy,* 55, 374-385.

Merrifield, M.S., McClintock, W., Burt, C., Fox, E., Serpa, P., Steinback, C., & Gleason, M. (2013). MarineMap: A web-based platform for collaborative marine protected area planning. *Ocean & Coastal Management* 74, 67-76.

McIntosh, B.S., Ascough Ii, J.C., Twery, M., Chew, J., Elmahdi, A., Haase, D., Harou, J.J., Hepting, D., Cuddy, S., Jakeman, A.J., Chen, S., Kassahun, A., Lautenbach, S., Matthews, K., Merritt, W., Quinn, N.W.T., Rodriguez-Roda, I., Sieber, S., Stavenga, M., Sulis, A., Ticehurst, J., Volk, M., Wrobel, M., van Delden, H., El-Sawah, S., Rizzoli, A., & Voinov, A. (2011). Environmental decision support systems (EDSS) development – Challenges and best practices. *Environmental Modelling & Software* 26, 1389-1402.

McIntosh, B.S., Seaton, R.A.F., & Jeffrey, P. (2007). Tools to think with? Towards understanding the use of computer-based support tools in policy relevant research. *Environmental Modelling & Software* 22, 640-648.

Perez-Soba, M., Maas, R. (2015). Scenarios: tools for coping with complexity and future uncertainty, in: Jordan, A.J., Tumpenny, J.R. (Eds.): The tools of policy formulation: actors, capacities, venues and effects. Edward Elgar Publishing Limited, Cheltenham, pp. 52-75.

Peterson, R.A., Merunka, D.R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research* 67, 1035-1041.

Peterson, R. A. (2001). On the Use of College Students in Social Science Research: Insights from a Second-Order Meta-analysis. *Journal of Consumer Research* 28 (3): 450-61.

Poch, M., Comas, J., Rodríguez-Roda, I., Sànchez-Marrè, M., & Cortés, U. (2004). Designing and building real environmental decision support systems. *Environmental Modelling & Software*, 19, 857-873.

Rodela, R., Reinecke, S., Bregt, A., Kilham, E., & Lapeyre, R. (2015). Challenges to and opportunities for biodiversity science–policy interfaces. *Environmental Science & Policy* 54, 483-486.

Russo, P., Lanzilotti, R., Costabile, M.F., Pettit, C.J., 2018. Towards satisfying practitioners in using Planning Support Systems. Computers, Environment and Urban Systems 67, 9-20.

Sambamurthy, V., Chin, W.W. (1994). The Effects of Group Attitudes Toward Alternative GDSS Designs on the Decision-making Performance of Computer-Supported Groups. *Decision Sciences* 25, 215-241.

Vahidov, R., Kersten, G., & Saade, R. (2014). An experimental study of software agent negotiations with humans. *Decision Support Systems* 66, 135-145.

Van Bavel, J.J., Mende-Siedlecki, P., Brady, W.J., & Reinero, D.A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences* 113, 6454-6459.

Van der Wal, M.M., de Kraker, J., Kroeze, C., Kirschner, P.A., & Valkering, P. (2016). Can computer models be used for social learning? A serious game in water management. Environmental Modelling & Software 75, 119-132.

Uran, O., Janssen, R. (2003). Why are spatial decision support systems not used? Some experiences from the Netherlands. *Computers, Environment and Urban Systems* 27, 511-526

Zapata, M.A., 2009. Deliberating across differences: Planning futures in cross-cultural spaces. Policy and Society 28, 197-209.