Facilitating Ontology Reuse with a Topic-Specific Trust Open Rating System

Holger Lewen

Institute AIFB, Universität Karlsruhe (TH), Germany
lewen@aifb.uni-karlsruhe.de

Abstract. In the experiment described in this report, we wanted to examine a) *how integration of tools for ontology reuse within the ontology engineering environment* and b) *quality information on the ontologies to be reused*, will facilitate the reuse of ontological ressources. In order to test that we conducted an experiment in which 20 researchers from the semantic web community participated. We have identified how tools can help during the steps of finding and integrating ontological content to reuse, and show that having quality information about the ontologies available facilitates the selection step for the participants of our study.

1 Introduction

In the experiment described here, we wanted to examine a) *how integration of tools for ontology reuse within the ontology engineering environment* and b) *quality information on the ontologies to be reused*, will facilitate the reuse of ontological ressources. In order to test that we conducted an experiment in which 20 researchers from the semantic web community participated. The participants came from 6 different institutions, most of them were not members of the NeOn project. They were assigned to one of three different experiment groups, each with different tools at their disposal. All participants were using the NeOn Toolkit[1] as ontology engineering environment, and had an Internet Browser at their disposal. The task was to extend an ontology solely consisting of the concept "Fish" to an ontology representing the fish domain, reusing ontological content found on the web. In order not to influence the results too much, it was left up to the user to decide for which concrete use case the ontology was modeled (i.e. the user could decide on which aspects or parts of the fish domain to focus, and how to model it). In order to to distinguish the effects of having an integrated tool, and having an integrated tool with quality information, we decided to have one control group (without integrated tool-support), one test group using the Watson plugin for NeOn Toolkit [1] and one test group using the Cupboard plugin. The Cupboard plugin is an extension of the Watson plugin also offering quality information on the ontologies retrieved from the Cupboard system[2]. In the following, we will first explain the experiment setup, then provide and analyse the experiment results and finish with a conclusion.

[1] http://www.neon-toolkit.org/

2 Experiment Setup

In this section, we will describe the experiment setup and its execution.

2.1 Goal of the Experiment

As mentioned before, the goal of the experiment was to evaluate empirically how ontology reuse can be facilitated by both tool integration and quality information on the ontologies. Facilitating ontology reuse means both allowing the user to reuse content more easily (offering help during the different steps of the reuse process), and also to produce better quality in a shorter time-frame. The time aspect is important, since in theory, given enough time, a user could review all candidate ontologies and import the best statements by hand. This is one of the reasons why we imposed a 20 minute time limit on all of the groups.

2.2 Technology Used in the Experiment

We will now briefly explain which technology was used in the experiment and what functionality it provides.

NeOn Toolkit The NeOn Toolkit[2] is an ontology engineering environment which is based on Eclipse and allows an easy integration of plugins through the Eclipse Plugin Framework.

Internet Browser The users were given access to the Internet and a list of semantic web search engines[3] which was set as the homepage of the browser.

Watson Plugin The Watson plugin for the NeOn Toolkit[4] [1] allows the user to directly access the Watson Semantic Web gateway[5] [**?**] from within the NeOn Toolkit. The user can right-click a concept in the ontology, and trigger the search from the context menu. Then, on the right hand side the resulting list of ontologies is displayed, and by clicking on an item in the list, the statements from that ontology containing the search-term are visible. If wanted, the user can create these statements automatically in the ontology by clicking on the add-button next to the statements (see also fig. 1). Note, however, that the ranking is not based on any quality information, so the user has to take care of assessing the quality without external help.

[2] `http://www.neon-toolkit.org/`
[3] `http://esw.w3.org/topic/TaskForces/CommunityProjects/LinkingOpenData/SemanticWebSearchEngines`
[4] `http://watson.kmi.open.ac.uk/editor_plugins.html`
[5] `http://watson.kmi.open.ac.uk/WatsonWUI/`

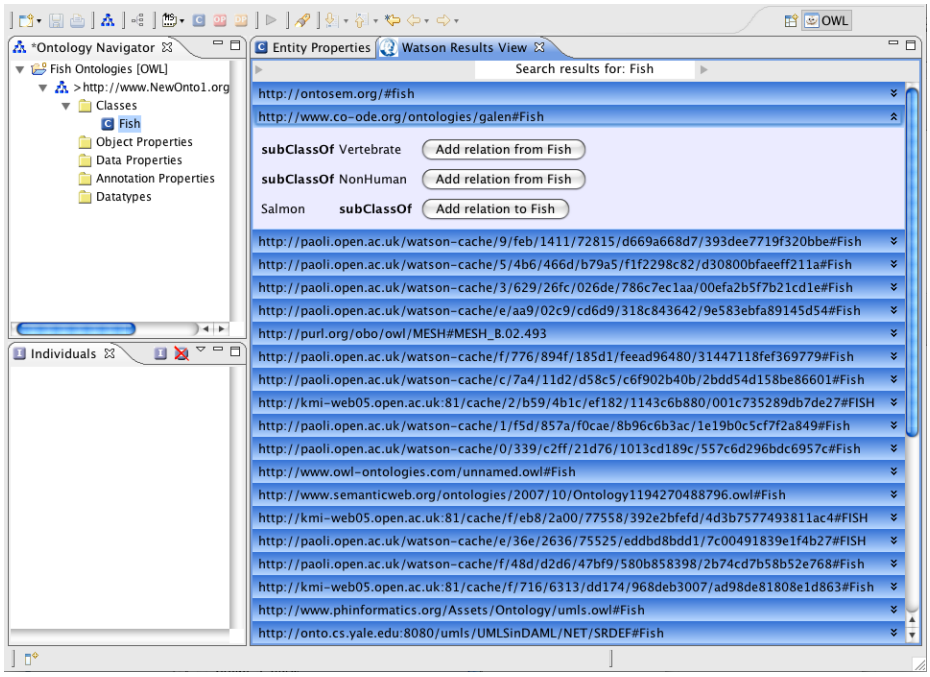

Fig. 1. This screenshot shows the Watson plugin displaying results within the NeOn Toolkit

Cupboard Plugin The Cupboard plugin for the NeOn Toolkit[6] is an extension of the Watson plugin, but connects to a specified ontology space within Cupboard [7]. We have loaded the ontologies space with ontologies we could find on the web and in Watson, as well as with some other ontologies we created. The ontologies were then reviewed by members of the NeOn Project. A special feature of the Cupboard plugin is that it can retrieve the overall ratings for each of the ontologies in the result-set, and rank the results accordingly (see fig. 2). During the course of the experiment, we relied on global trust ratings, that means the identity of the user was unknown to the system, and thus all users were presented with the same ranking order. Another feature we added to the Cupboard plugin was the ability to add multiple superclasses (all superclasses from found concept to the root) and subclasses (all classes below the concept) with one-click. These are currently not available in the Watson plugin, since without knowing anything about the quality of the ontology, we felt that it would be too risky to allow adding too many statements blindly. Since the user is presented with quality information on the ontology in the Cupboard plugin, we felt this feature would make sense.

While it would have been possible in the Topic-Specific Open Rating System (TS-ORS)[3, 4, 2] to review single statements within the ontology, we felt it would make no sense to ask people to review at this level of granularity. Many of the properties can only be evaluated meaningfully if the ontology is assessed as a whole, and not on the statement level. So basically the ranking within the Cupboard plugin is based on reviews on the complete ontology, since the found statements are part of these ontologies.

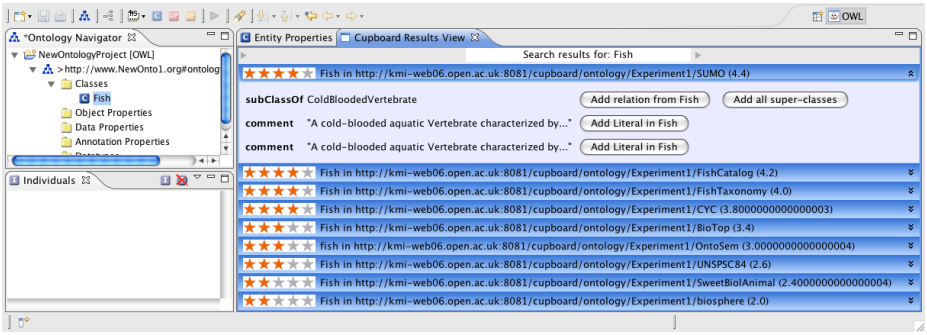

Fig. 2. This screenshot shows the Cupboard plugin displaying results within the NeOn Toolkit

[6] Which you can obtain from `http://www.aifb.uni-karlsruhe.de/WBS/hle/experiments` (User is "reviewer", Password is "iswc2009")

[7] You can check out the ontology space here: `http://kmi-web06.open.ac.uk:8081/Cupboard/Experiment1`

2.3 Tasks to be Executed

Each user was given a task description (see fig. 22) explaining the task to be executed. The task was to extend an ontology containing only the concept "Fish" with ontological content found on the web, both with superclasses (ideally aligning it to an upper level ontology) and with subclasses. A facilitator was present at all times to provide individual help and answer questions related to handling the tools. Questions potentially affecting the outcome of the study (e.g. which ontology should I reuse) were not answered. The participants were assigned to one of three groups, each having different tool-support at their disposal. For each group, there was an adapted version of the NeOn ontology reuse methodology, mentioning tools (if available for that group) that could help in each step. All participants had at most 20 minutes to complete the task, but could stop earlier if they felt the ontology they created was satisfactory for them. We will now explain the distinctions between the three groups.

Group 1 The first group was the control group, which was not allowed to use the Watson- or Cupboard plugin. The participants therefore had to create the statements they wanted to reuse manually within the NeOn Toolkit. They also had to search for content to reuse using the web browser. The methodological guidelines (see fig. 23) provided a sort of best practice for ontology statement reuse.

Group 2 The second group had all the means of the first group, but in addition the Watson plugin. The plugin allowed the users to search Watson directly from the NeOn Toolkit and also to import statements by the click of a button (see fig. 2). Only if content outside Watson was added, the user had to do this by hand. The methodological guidelines (see fig. 24) were adapted to mention the Watson plugin in the steps that could benefit from it.

Group 3 The third group had access to all means of the first group, and in addition to the Cupboard plugin. The Watson plugin was not available. The Cupboard plugin allows to search for ontologies within Cupboard, and to import statements (or multiple statements) with one click. Furthermore, an overall rating is displayed for the ontologies based on reviews written within Cupboard. The ontologies are ordered based on that rating (see fig. 2). In the alpha version of the plugin used for the experiment, advanced features as local trust (user-specific trust) or displaying the reviews from within the plugin were not available. The methodological guidelines (see fig. 25) were adapted to mention the Cupboard plugin in the steps that could benefit from it.

2.4 Role of the Facilitator

The facilitator was given guidelines (see figures 26 and 27), explaining how to conduct the experiment. The facilitator's role was to find participants and assign them to one of the three groups. Furthermore, the facilitator handed out

a printed version of the experiment's task and of the methodological guidelines. The hardware was also set up by the facilitator. The facilitator was responsible to take care of screen-capturing and saving the ontologies. During the experiment, questions regarding the handling of the tools were also answered by the facilitator. After the experiment ended, the facilitator gave a short demonstration of the Cupboard plugin to participants from group 1 and 2. Ultimately, the facilitator was also responsible for sending the questionnaire to the participants.

2.5 Running the Experiment

When a participant was found and assigned to one of the three groups, the task description and the methodological guidelines for that group were handed out. After the participant had read the two documents, he or she had to create an ontology with the concept "Fish" using the NeOn Toolkit. The creation of the ontology and concept was not part of the experiment, but rather an initial step to ensure the user could work with the NeOn Toolkit (e.g. create classes). Once the user indicated he or she was ready, the screen-capturing was started, and the time was measured. After 20 minutes (users could stop earlier if they felt the task was completed) the ontology was saved and the screen-capturing stopped. Users from group 1 or 2 were given a quick demonstration of the Cupboard plugin after they finished their task. Then the questionnaire was sent to the participant to gather feedback.

2.6 Questionnaire

The questionnaire was sent to all participants with the purpose of finding out more about the participants (e.g. their level of expertise, understanding of the task) and gathering feedback as well as impressions. The questions can be found in figures 28, 29, and 30. Because the answers are all subjective statements, they cannot be used alone to assess the quality of a tool. But they provide insight into whether users like the tool or not, if they understood the task presented, and how they would judge their expertise in different areas. We therefore have them in addition to the videos and ontologies which are the outcome of the modeling, and can be assessed independently of the user's subjective impression.

2.7 Preparation of Cupboard for the Experiment

Since gathering reviews for all ontologies in Watson was unrealistic in the time given for preparing the experiment, we decided to focus on a limited subset of Watson mentioning fish, and asking qualified members of the NeOn project to review the ontologies. We also added a few extra ontologies not found in Watson or elsewhere on the Web (bad and good ones) to see a) if the reviewing worked (if the bad ontology would be given a low rating by the reviewers), and b) if people indeed would rely on the rating information displayed by the system to make their selection choice. The ontologies and the Cupboard Space can

be found here: `http://kmi-web06.open.ac.uk:8081/Cupboard/Experiment1`. We thought this would be a realistic setup for a Cupboard Space, because the idea is that users also add their own ontologies, which cannot be found elsewhere on the web. We then asked ontology engineers within NeOn to review the ontologies in the ontology space based on 5 properties: Reusability, Correctness, Complexity, Domain Coverage and Modelling (these are basically the properties mentioned in [5]). The reviewers could also add trust or distrust statements to other reviews. For the experiment we decided to use the average over the five properties based on global trust as overall rating for each ontology, so the ranking order and results would be the same for all users. At a later stage, the plugin will also allow the user to assign weights to different properties and to identify and use ratings based on local trust.

2.8 Hardware and Software Used

The experiments were conducted on a 15.4"MacBook Pro running MacOS X Leopard. All groups were using the latest available extended version of the NeOn Toolkit (Version 1.2.2 B904 extended for Mac). Group 2 was using a version of the toolkit with the Watson plugin installed, Group 3 was using a version of the toolkit with the Cupboard plugin installed.

3 Results

In this section we will simply present the results of the experiment, with the analysis following in the next section. We will both provide a table displaying information about the ontologies produced, and the results of the questionnaires filled out by the participants after the experiment. In order to keep the evaluation transparent, the interested reader can access the videos, ontologies and questionnaires online[8] to draw his or her own conclusions from the raw data.

3.1 Ontologies

In order to somehow quantify the ontologies produced by the different groups, we took a look at both the number of axioms and the quality of the ontology with respect to good engineering practices. Figure 3 presents a table with information on the ontologies engineered during the experiment. Since there was a bug regarding wrong use of namespaces in both the initial version of the Watson and Cupboard plugin, we have corrected the created ontologies manually where necessary, so they can be opened by all ontology engineering environments. The bug was fixed and the later experiments were conducted with the error-free version. The bug should not have affected the outcome of the experiment.

3.2 Results Questionnaire

In order to gather demographic data about our users and get feedback from them, we have asked each participant to fill out a questionnaire after the experiment was finished. As can be seen in figures 28, 29, and 30, some questions are group specific (only for participants in group 3, or only for those not in group 3). We have created charts based on the answers to each of the questions comprising the results of all groups side by side, so it can be easily seen how the different groups compare to each other. The charts can be found in figures 4 trough 21.

4 Analysis

In order to draw conclusions from the experiment, we analysed both the questionnaire and the ontologies produced. We will start by giving some basic information about our participants. We had a total of 20 participants from 6 different academic institutions. Most of the participants were PhD students, but we also had 2 postdocs and 1 professor participating. We selected people both from within the NeOn project (4 participants) and outside the NeOn project (16 participants). The idea was to have a heterogeneous group and not to be biased in the selection of participants. As can be seen in fig. 4, more than half of the total participants gauged their experience with the NeOn Toolkit and other tools as

[8] `http://www.aifb.uni-karlsruhe.de/WBS/hle/experiments` (User is "reviewer", Password is "iswc2009")

Group	N° in Group	Class Count	Object Properties Count	SubClass Count	Equivalent Class Count	Different Sources Used	Self-Created Axioms
1	1	7	3	5	0	5	2
1	2	6	0	2	1	4	1
1	3	10	1	8	0	2	7
1	4	14	1	12	0	1	12
1	5	5	0	2	2	3	0
1	6	11	0	8	2	3	0
2	1	49	1	45	0	7	6
2	2	141	0	140	0	10	0
2	3	16	1	13	0	2	3
2	4	131	0	132	0	4	0
2	5	76	0	75	0	8	0
2	6	24	0	23	0	14	0
2	7	20	0	19	0	5	0
3	1	1429	0	1431	0	3	0
3	2	581	0	582	0	3	0
3	3	591	0	593	0	2	0
3	4	863	0	877	0	3	0
3	5	591	0	593	0	2	0
3	6	592	0	594	0	2	1
3	7	591	0	593	0	3	0

Fig. 3. A Table Comprising Quantitative Information on the Ontologies Created in the Experiment. Different sources used refers to the number of ontologies from which statements were reused. Self-Created Axioms refers to axioms which were not found in another ontology, but created from scratch or based on knowledge acquired from other non-ontological sources.

beginner. Also more than half of the total participants (see fig. 5) thought of themselves as beginners with regard to ontology engineering. Nevertheless, all the vast majority of the users had no trouble understanding the task (see fig. 7), regardless of whether they were ontology engineering experts or beginners. We will now analyze the results group by group and then relate the different groups to each other.

4.1 Group 1

Since group 1 is our control group without tool support, it is both supposed to serve as a baseline and to gather insight into how people can benefit from all the ontology search engines on the web, and how easy it is to reuse ontologies at the moment. Because this was the group for which we thought it would be hardest to complete the task in 20 minutes, we made sure that the expert ontology engineers would be part of this group (see fig. 5). We thought it would be only fair to have a rather high baseline, and not be accused of only putting novice users in the control group to tweak the outcome. Of the six participants of the group, four are (and judge themselves) ontology engineering experts, one is a beginner and one has moderate ontology engineering experience.

Insights Gained from the Videos While analysing the video, it was surprising to see that even expert users had trouble finding ontological statements to reuse and to integrate them into their ontology in the NeOn Toolkit. Most of the results produced by the ontology search engines were confusing to the participants since they had no clear ranking. Also most of the tools had issues with usability, leading to situations were the participants expected a certain action and triggered another. This can be seen in the videos when a user clicks somewhere, only to immediately go back and click at another button until the desired outcome is reached. Most of the participants gave up searching for content to reuse, and started modeling the ontology directly from within the toolkit. Some where using their own knowledge about the fish domain, some were consulting Google[9] or Wikipedia[10]. So in a sense, most participants did give up on the idea of reusing existing knowledge after trying at the beginning and reverted back to creating the content from scratch.

Insights Gained from the Questionnaire The observation from the video that the participants had trouble finding and integrating statements to reuse was confirmed by the answers from the questionnaire. All of the participants in group 1 said they had trouble finding statements to reuse (see fig. 11) and all but one had trouble integrating found statements (see fig. refQ10). Also, all but one participant had problems selecting the ontology statement to reuse from the statements found during the search (see fig. 12). Regarding the methodology

[9] `http://www.google.com/search?hl=en&q=fish&btnG=Suche&meta=`
[10] `http://en.wikipedia.org/wiki/Fish`

presented, half of the participants found it at least reasonably useful, with the other half not giving any information. Since group 1 contained mainly expert users, these might already have their own methodology internalized, and not rely on the one presented in the experiment (see fig. 14).

Insights Gained from Analysing the Ontologies When checking the metrics for ontologies produced by group 1 (see fig. 3), it becomes evident that all of the resulting ontologies are small in size (ranging from 7 to 14 classes). While all tried to reuse at least some ontology statements they found on the web, the majority started to add their own axioms at some point in the experiment (see the see self-created axioms column). One thing that was interesting to see is that half of the participants linked the reused classes to their original location by either creating equivalent classes with the reused URI, or creating the new class directly with the reused URI. One problem that we also observed mainly with participants in group 2 was the reuse of superclasses from different (upper level-)ontologies. This led to the result that in some ontologies "Fish" was both a subclass of "Vertebrate" (this statement comes from CYC) and "ColdBlooded-Vertebrate" (this statement comes from SUMO), with "ColdBloodedVertebrate" being a subclass of "Vertebrate" (coming from SUMO). Here the participant should have corrected the taxonomy by removing the superclass "Vertebrate" from "Fish". Other ontologies contained only references to the concept "Fish" in other ontologies. For most of the ontologies created, the purpose was not clear and the requirements from the task description were not fulfilled. It was evident that all participants had problems both finding content to reuse and also assessing and integrating found content into their ontology. This is why most participants started modeling without reusing ontological content. We invite the reader to access the produced ontologies[11] themselves to judge their quality. It is still important to note that some of the axioms found in these ontologies were not reused (it can be seen in the videos which were created by hand and which were found and then reused).

4.2 Group 2

More than half of the users in group 2 judged themselves beginners with respect to the NeOn Toolkit and Watson plugin (see fig. 4), but most of them have moderate ontology engineering experience (see fig. 5).

Insights Gained from the Videos While the user interface of the Watson plugin is simple and straight forward to use (see fig. 1), most participants had problems selecting useful statements to reuse from the list of results. They browsed the list of results, and then started adding statements from all over the list. It seems that many participants did not take the time to actually assess

[11] `http://www.aifb.uni-karlsruhe.de/WBS/hle/experiments` (User is "reviewer", Password is "iswc2009")

the found statements and check whether they integrate well into the current ontology (e.g. the statement serves a similar purpose in the original ontology). They rather started adding statements from as many sources as possible. Also people started to look for statements related to reused statements, not necessarily focusing on the fish domain. In two cases, the participants started to create classes themselves, not using the plugin.

Insights Gained from the Questionnaire As we expected after analysing the videos, all but 1 participant had no trouble finding potential statements to reuse (see fig. 11), but most had trouble selecting the ontology statement from the list of found statements (see fig. 12). While triggering the search (and thereby producing the results) from within the NeOn Toolkit is easy, selecting the useful statements from this list is not. Since the ranking of results in Watson (and therefore also in the Watson plugin) is mainly based on the Lucene[12] engine, the quality of the ontologies is not factored into the ranking. Therefore a user had to look at all found statements before knowing which were good and which not. Integration was then easy (see fig. 13), since it was done automatically with the click of a single button. All users found the provided methodology at least reasonably useful (see fig. 14). Note that the methodology was adapted for group 2 to explain in which steps the Watson plugin can be used.

Insights Gained from Analysing the Ontologies First of all, all ontologies produced by group 2 (see fig. 3), are larger in size than those produced by group 1. Also, on average, twice as many different sources were used by participants of group 2 compared to group 1 (7.1 vs. 3 on average). These two metrics can easily be explained by the use of the Watson plugin. By the nature of the plugin, many results are presented from which statements can then easily be reused. The problem with the resulting ontologies is that users often blindly reused statements without checking whether they need them in their ontology. Most ontologies found on the web were build for a special purpose, so the way the world or a domain is modeled varies based on the requirements. Sometimes the taxonomy is very fine grained, while in other cases only relevant information is included. When reusing blindly from too many sources, the resulting ontologies face quality problems. In one ontology for example, "Fish" is a subclass of "AnimalFoods", "AquaticOrganism" and "Seafood". In another ontology "Fish" is a subclass of "AquaticOrganism", "MarineAnimal", "Organic", "Seafood" and "Vertebrate". Normally one would expect that these superclasses would themselves be in some sort of hierarchy, if they were at all needed within one ontology. As said before, there are good reasons to have each of them as a superclass in one ontology, depending on the purpose of the ontology, but when combined, the hierarchy does not make much sense anymore. In another ontology, which chose to use the ontology to describe fish dishes, "Fish" is a subclass of "NonHuman". It is unclear why this statement would be needed in this context. Most of the

[12] `http://lucene.apache.org/`

ontologies created face similar problems. The users reused content from various sources, but without assessing which statements they might need. So one can say that reusing is easy with the Watson plugin, but knowing what to reuse is still hard.

4.3 Group 3

Most of the users in group 3 judged themselves beginners with respect to the NeOn Toolkit and the Cupboard plugin (see fig. 4), but the majority of them has moderate ontology engineering experience (see fig. 5).

Insights Gained from the Videos Most of the users finished the task very quickly, reusing mostly two or more of the first four ontologies in the result list. The users also heavily used the "add all superclasses" and "add whole sub-branch" feature. This feature was added to the Cupboard plugin, since the ontologies in Cupboard can be reviewed, and the ratings seen from within the plugin. So we felt it was more secure if a user knew an ontology was rated 4 stars and then chose to add more content based on a single statement than in the case of the Watson plugin, where it is unsure what effect adding a sub-branch or all superclasses can have. Since the user can only see one axiom in both Cupboard and the Watson plugin, a certain trust should be placed in the ontology before blindly reusing statements. This was offered by the reviews from NeOn members, which also influenced the ranking of the results within the plugin.

Insights Gained from the Questionnaire The user's answers confirmed our impression from the videos that neither finding, nor selecting or integrating ontology statements posed a problem within the group. All but one participant said they did not have trouble finding ontology statements to reuse (see fig. 11), and none had trouble selecting it (see fig. 12), or integrating it (see fig. 13) into their ontology. All but one users found the provided methodology very useful (see fig. 14). Note that the methodology was adapted for group 3 to explain where the Cupboard plugin can be used.

Also the time needed to execute the task of the experiment was perceived low by all but one participant (see fig. 6), in contrast to participants from groups 1 and 2. Group 3 had some specific questions related to the functionality offered by the Cupboard plugin. All of the participants found the possibility to search for statements to reuse from within the NeOn Toolkit, and adding them (and potentially multiple super- and subclasses) directly very useful (see figs. 14, 17, and 18). All participants found the ranking of the statements based on the reviews at least reasonably useful, 3 found it very useful (see fig. 16). All but one participant said that the ranking helped in the selection process of the statements (see fig. 19). So the main claim, that the TS-ORS facilitates ontology reuse by providing quality information about the ontologies for the selection process was confirmed.

Insights Gained from Analysing the Ontologies All ontologies produced by group 3 (see fig. 3) are big in size, which is mainly due to the possibility of adding multiple statements with one click and the size and structure of the ontologies available in Cupboard. All participants reused at most 3 ontologies, mostly SUMO as upper level ontology, and two ontologies containing mostly information about different fish type and species. Except for one participant, all participants used only SUMO as upper level ontology, and not CYC. One user decided to use both. 2 users used the scientific classification of fish, one together with the information on fish types from another ontology, and only the scientific classification. The rest chose to only reuse information on the different fish type. In general only ontologies rated highly were reused by the participants, so there is no obvious quality problem with the resulting ontology. Because of the quality information displayed, they did not blindly add statements from all ontologies, but only from the best rated ones.

4.4 Relation Between Groups

As discussed in the group analysis before, one could say that group 1 had the hardest time finding, selecting and integrating ontology statements to reuse in their ontology. The participants did not manage to produce a suitable ontology satisfying the task requirements given the 20 minutes time limit. Participants of group 2 had no trouble finding and integrating ontology statements thanks to the Watson plugin, but did not know which statements to reuse. This often led to ontologies comprising statements from many different ontologies, resulting in ontologies having serious modeling issues. The last group completed the task very easily, based on the quality information provided through the Cupboard plugin. After the recording was stopped, all participants of group 1 and 2 were given a quick demonstration of the Cupboard plugin. In the questionnaire, they unanimously stated they thought the plugin was very helpful for the reuse task performed in the experiment (see fig. 21) and that they would have liked to use it during the experiment (see fig. 20).

4.5 Remarks on Linked Data

As stated before, some participants in group 1 manually took care of linking the ontology created the ontology they reused, be it by reusing the URI (which can be problematic) or creating equivalent classes. Both the Watson and the Cupboard plugin offer the functionality to automatically create equivalent classes for each statement that is imported. We have disabled this feature for the experiments, since the equivalent classes show up in the ontology (which results in the user seeing two classes with the same name, one of which is the local class and the other is the equivalent class with the URI), which can confuse the user and hinder usability. For future versions, we plan to automatically provide mapping files linking the created ontology to the ontologies from which statements were reused.

5 Conclusion

In the course of the experiment we could see what problems also experienced ontology engineers face when trying to reuse ontological content. The main three problems are finding the statements to reuse, then assessing them and lastly integrating them. The problem of finding statements to reuse is nowadays addressed by many ontology search engines[13]. However, most of them simply store whatever rdf, owl or foaf document they find on the web without prior quality checks. Also the interface is still confusing users and when statements have been found and selected, they often have to be entered by hand in the ontology engineering environment. All these problems could be seen when analysing the results of participants in group 1, which had no integrated tool support for search, selection or integration.

The Watson plugin addresses the problems of search and integration, since it uses the Watson API to expose the search functionality directly in the NeOn Toolkit as a plugin, and also allows for easy integration of found statements. But it does not offer quality information on the indexed ontologies, thus leaving the selection process entirely to the user. It could be seen in the experiment that users could create larger ontologies more easily with the Watson plugin, but still had trouble deciding which statements to reuse. One could say it is easy to reuse ontological content, but it is also easy to create an ontology that does not adhere to good ontology engineering practice.

Cupboard closes this gap by offering users the possibility to review ontologies and trusting reviewers. The TS-ORS computes ratings on each ontology based on weights for each property and the best reviews. The best reviews are determined by algorithms that produce both local trust (user-specific) and global trust (not user-specific). The Cupboard plugin uses these ratings to produce a ranking of the results, and also displays the overall rating as stars and numerically. This feature was welcomed by the users and facilitated the selection process. Overall the ranking must have been quite accurate, since the users only reused statements from one of the top four ranked ontologies, even though they checked the statements in the lower ranked ontologies as well. In comparison users of the Watson plugin reused statements from all over the result-list, since the list had no quality specific order.

Of course the quality of an ontologie resulting only from reuse can only be as good as the ontologies available to the search engine. And in most of the cases ontologies will deviate from existing ontologies, so they can not be created only reusing content. There are many cases, however, where an empty ontology can quickly be populated with existing axioms, which can then later be extended, moved, refactored or deleted. For our experiment in the fish domain, it was not too difficult finding ontologies with which we populated Cupboard or to find people reviewing them. We believe that taking the extra effort of reviewing the ontologies goes a long way in facilitating and encouraging reuse. For the partic-

[13] `http://esw.w3.org/topic/TaskForces/CommunityProjects/LinkingOpenData/SemanticWebSearchEngines`

ipants of group 3, building an initial ontology about fish which they could then later extend and reengineer was a matter of a few minutes. Compared to the results of the other two groups, we feel confident to say that we have shown that the Cupboard Plugin facilites reuse by solving all problems normally encountered by the users, namely finding, selecting and integrating existing ontological content. We have also shown that having the quality information and ranking of the TS-ORS available facilitated the reuse process for the participants of our study.

6 Appendix

6.1 Results Questionnaire

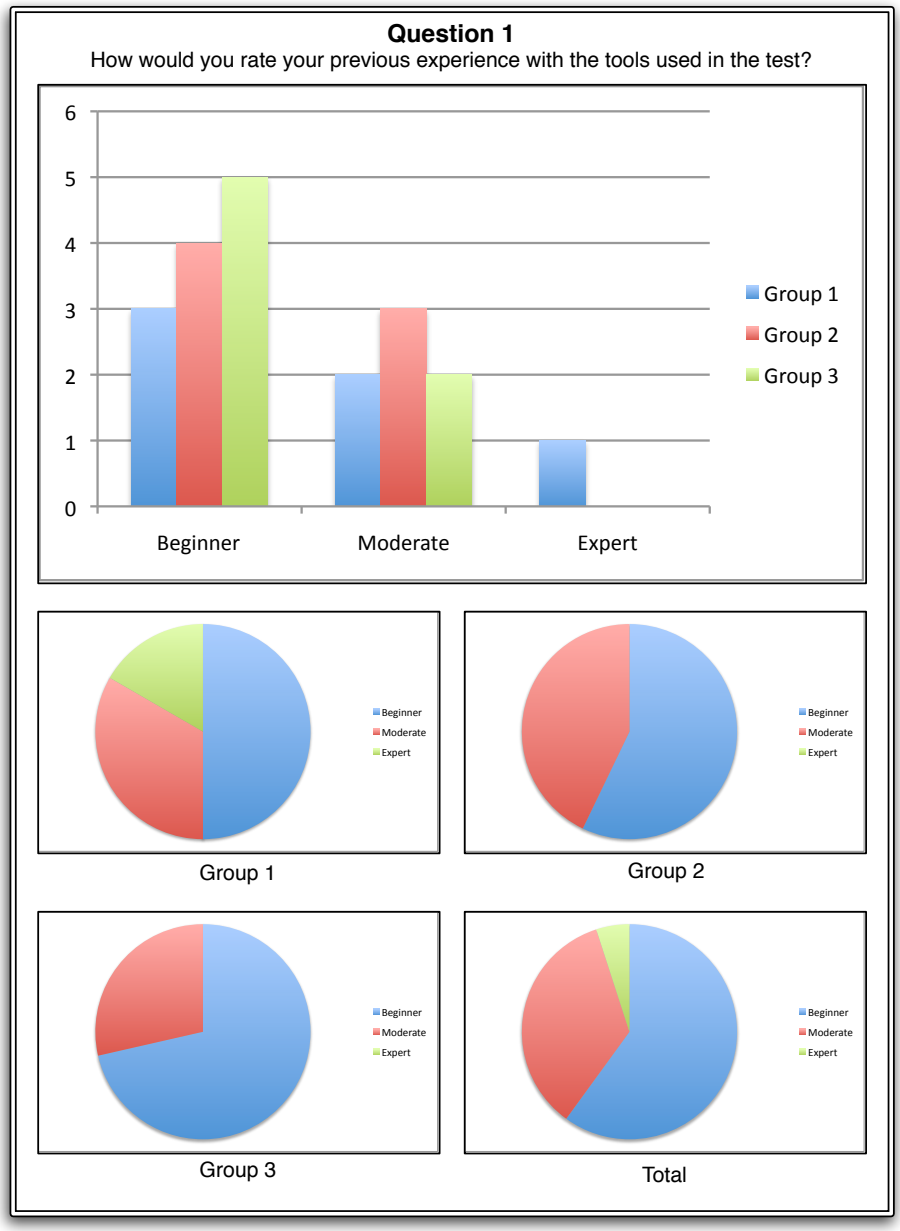

Fig. 4. A graphical representation of the results for question 1.

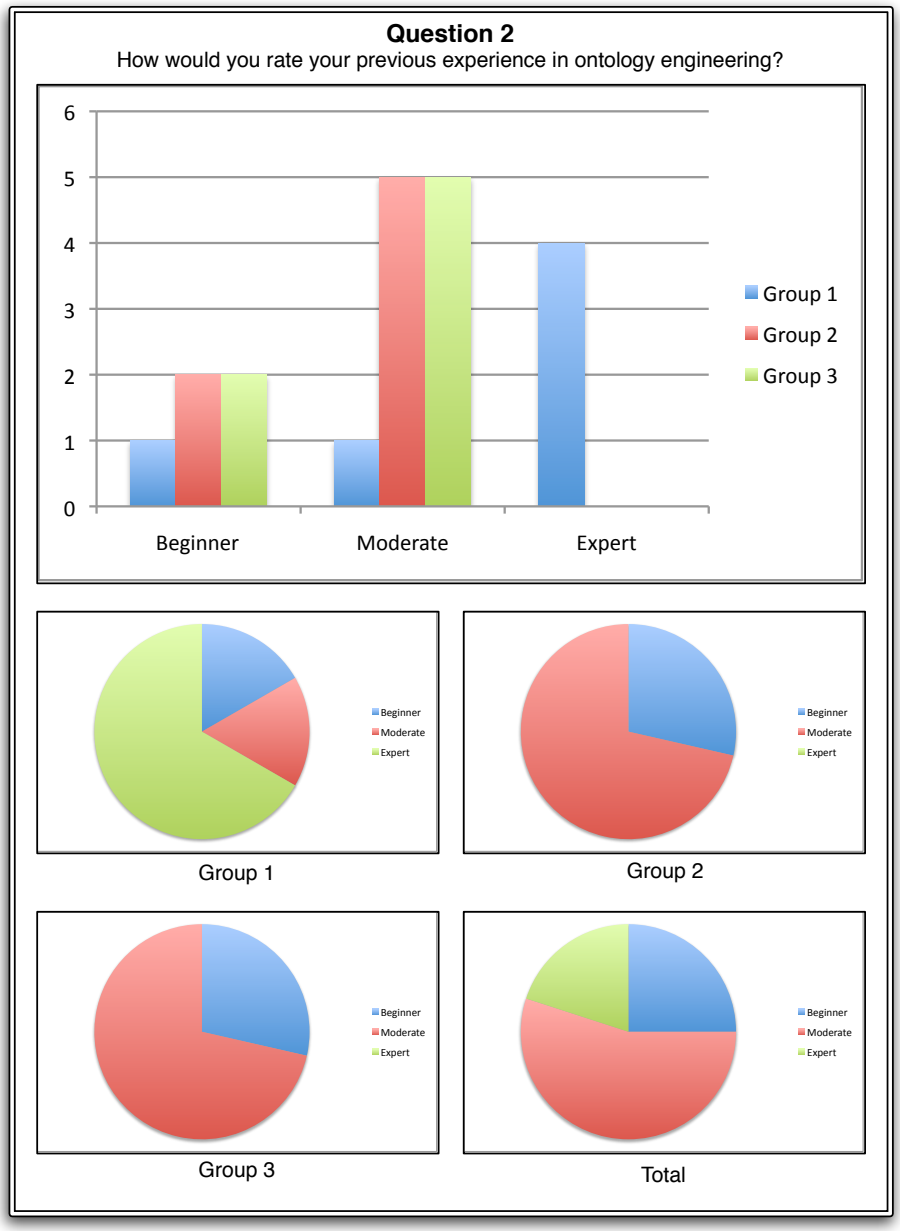

Fig. 5. A graphical representation of the results for question 2.

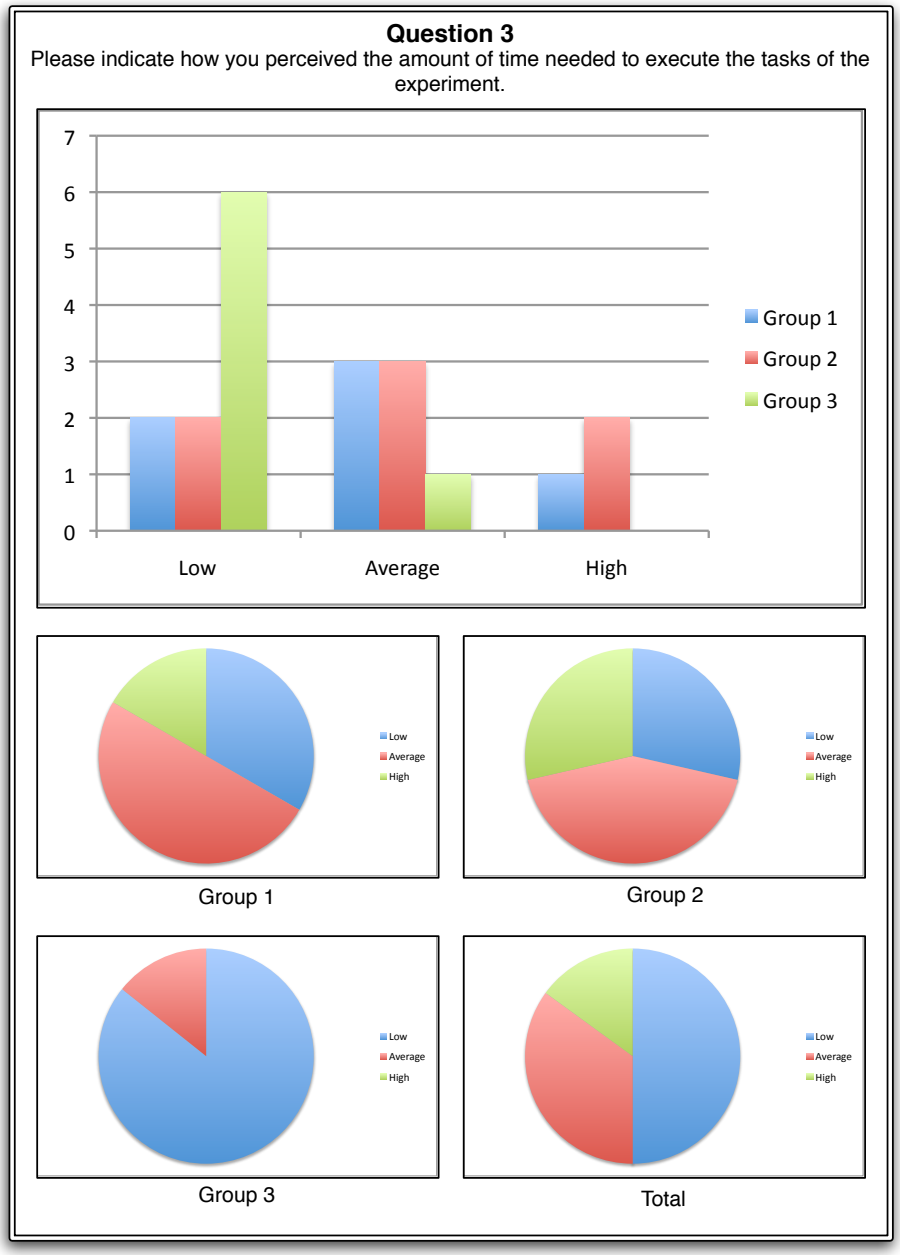

Fig. 6. A graphical representation of the results for question 3.

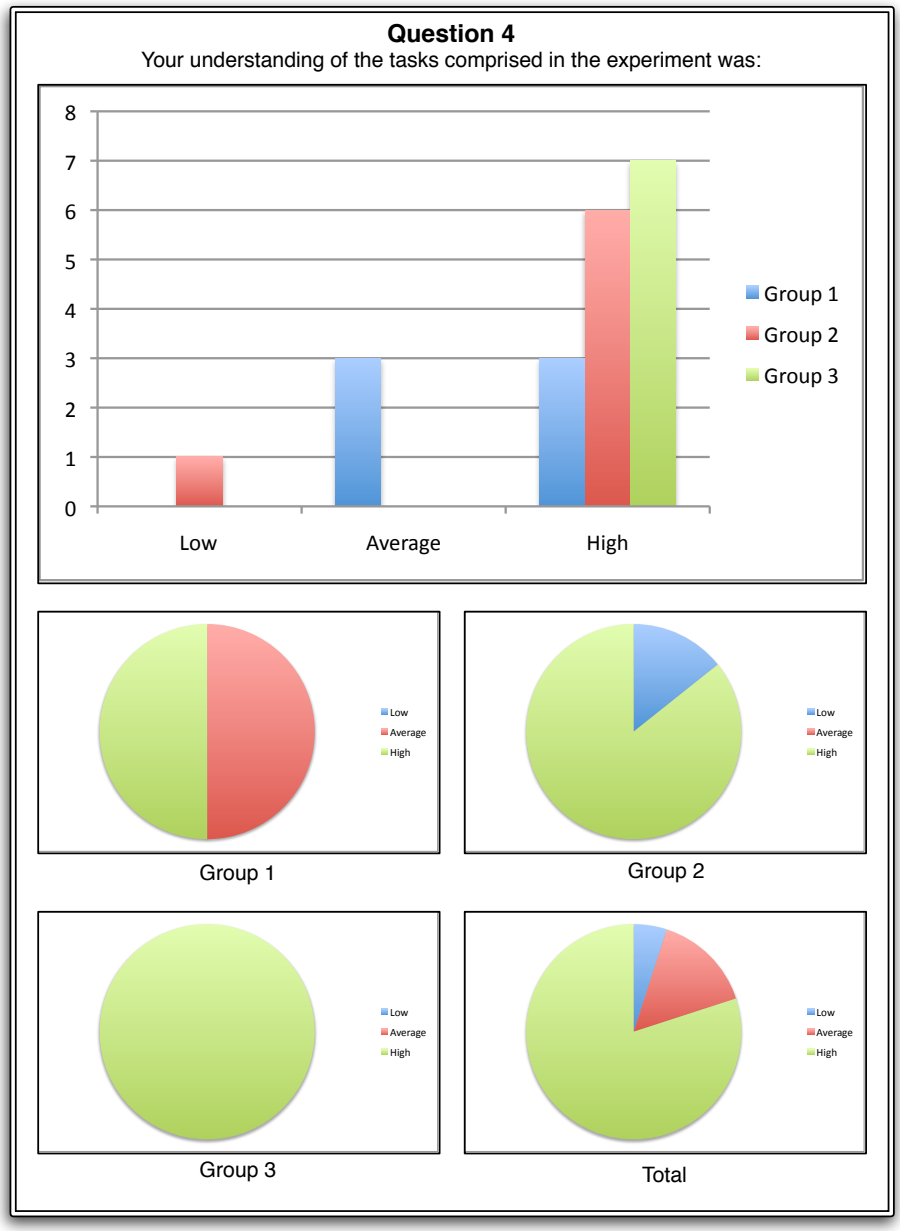

Fig. 7. A graphical representation of the results for question 4.

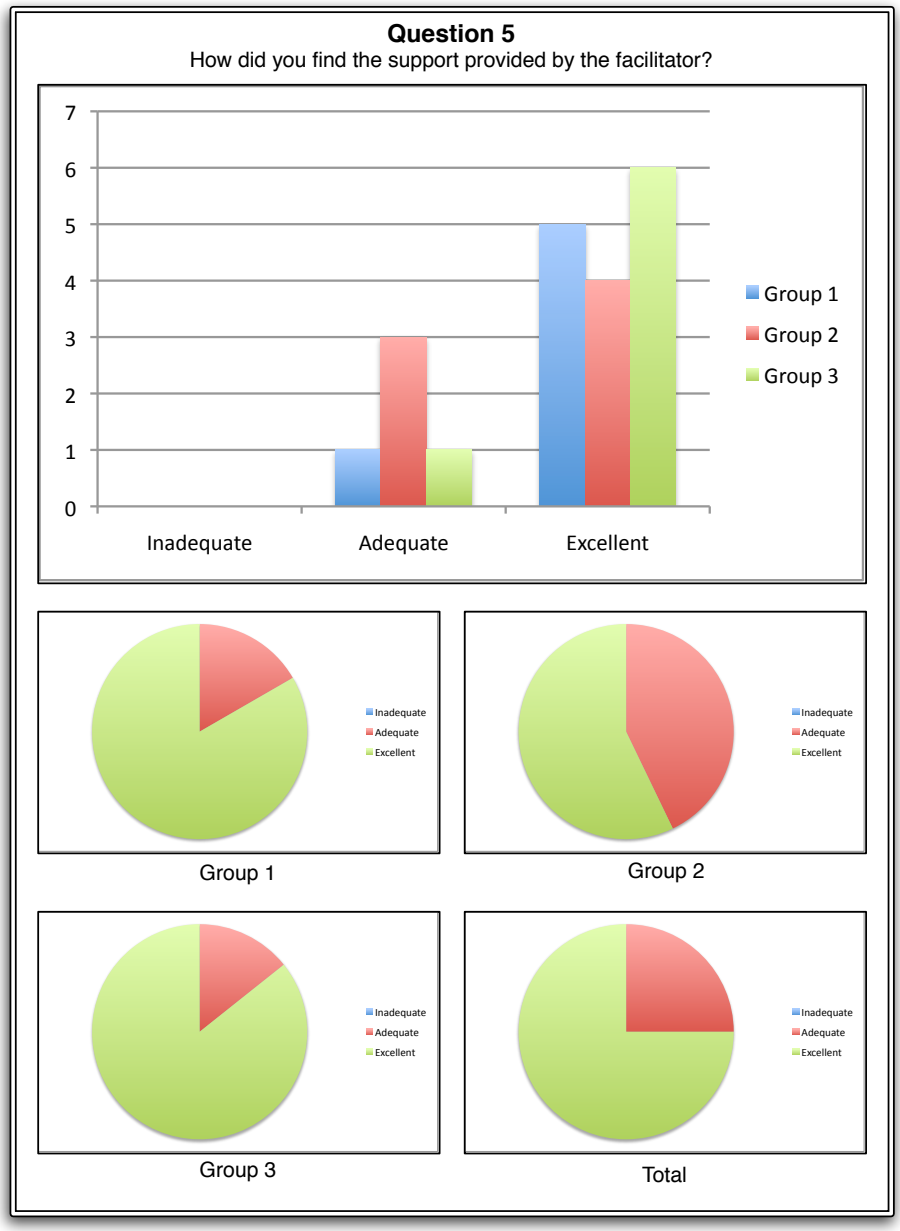

Fig. 8. A graphical representation of the results for question 5.

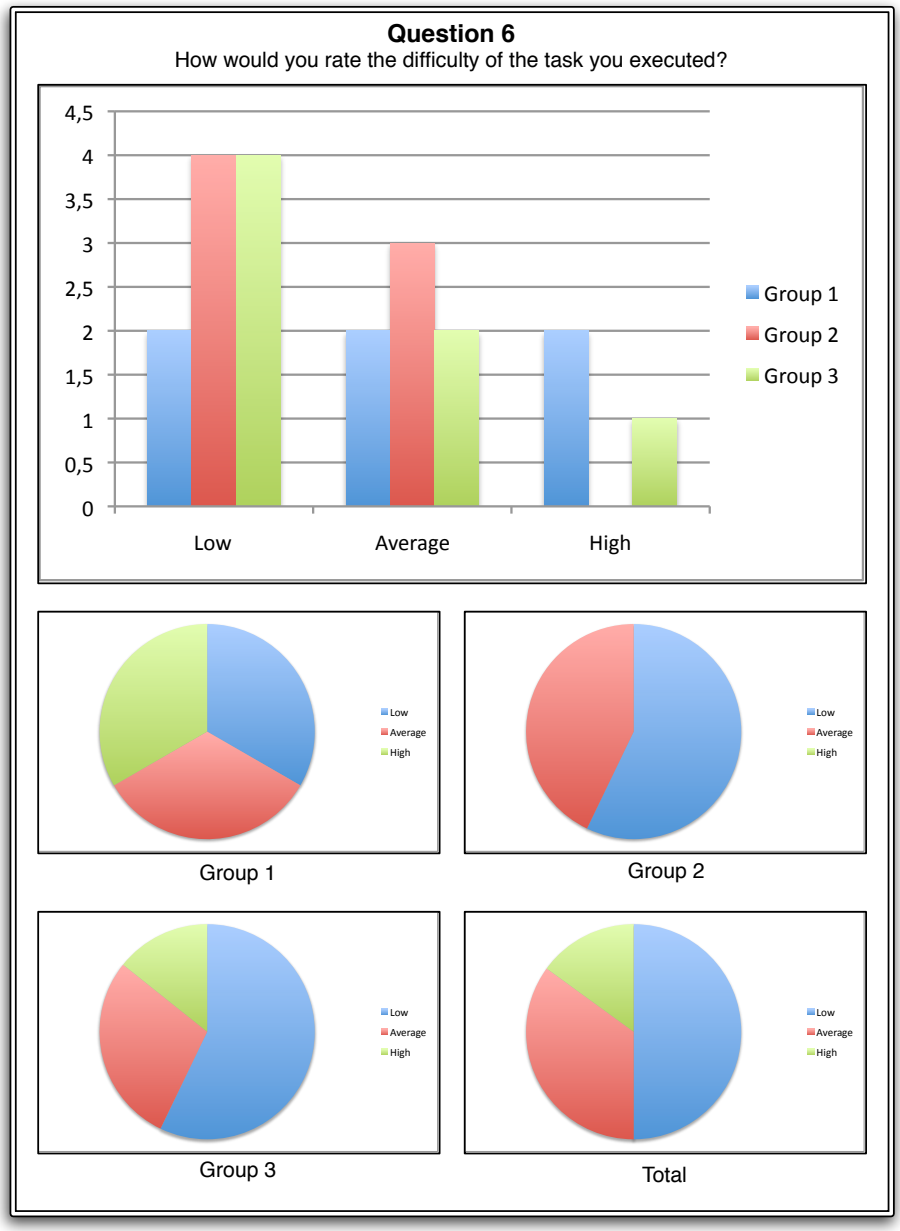

Fig. 9. A graphical representation of the results for question 6.

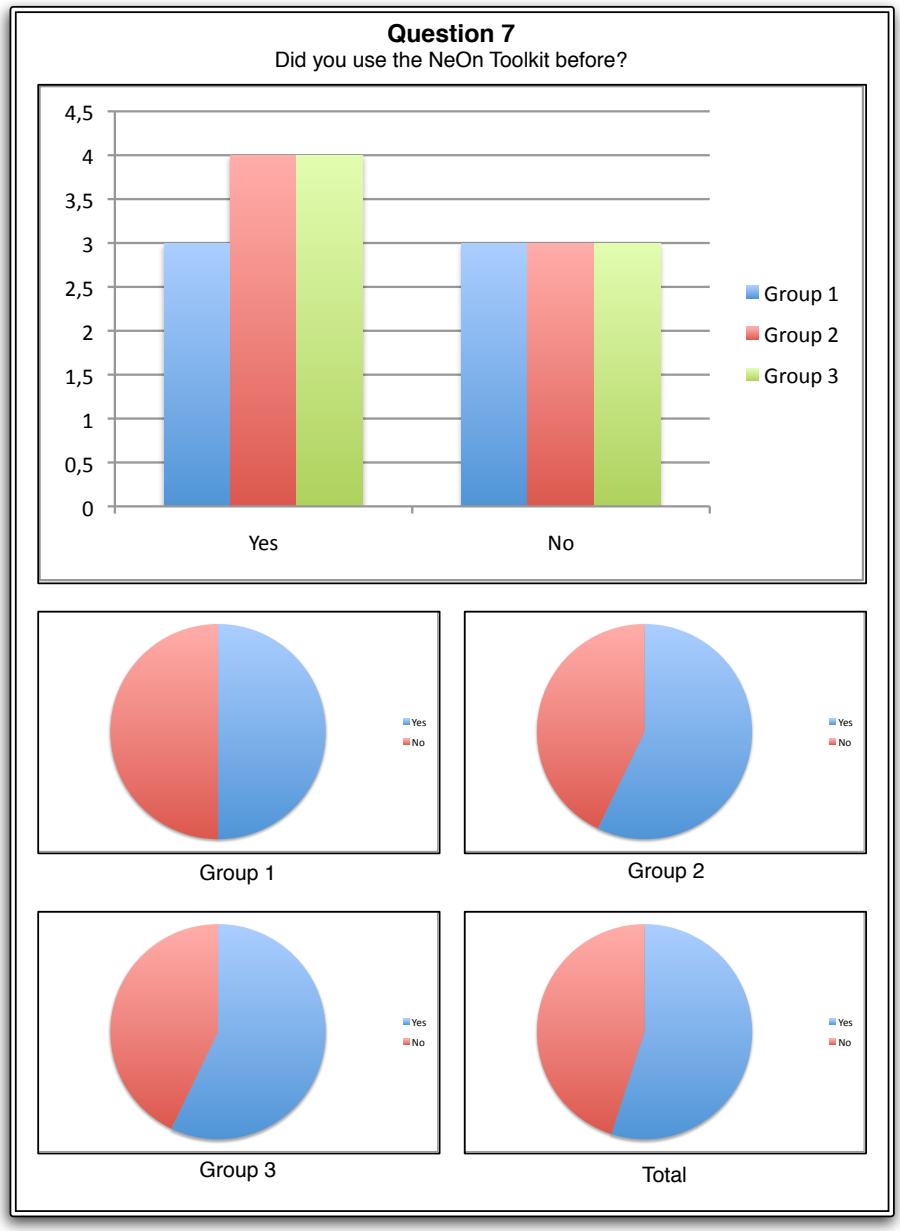

Fig. 10. A graphical representation of the results for question 7.

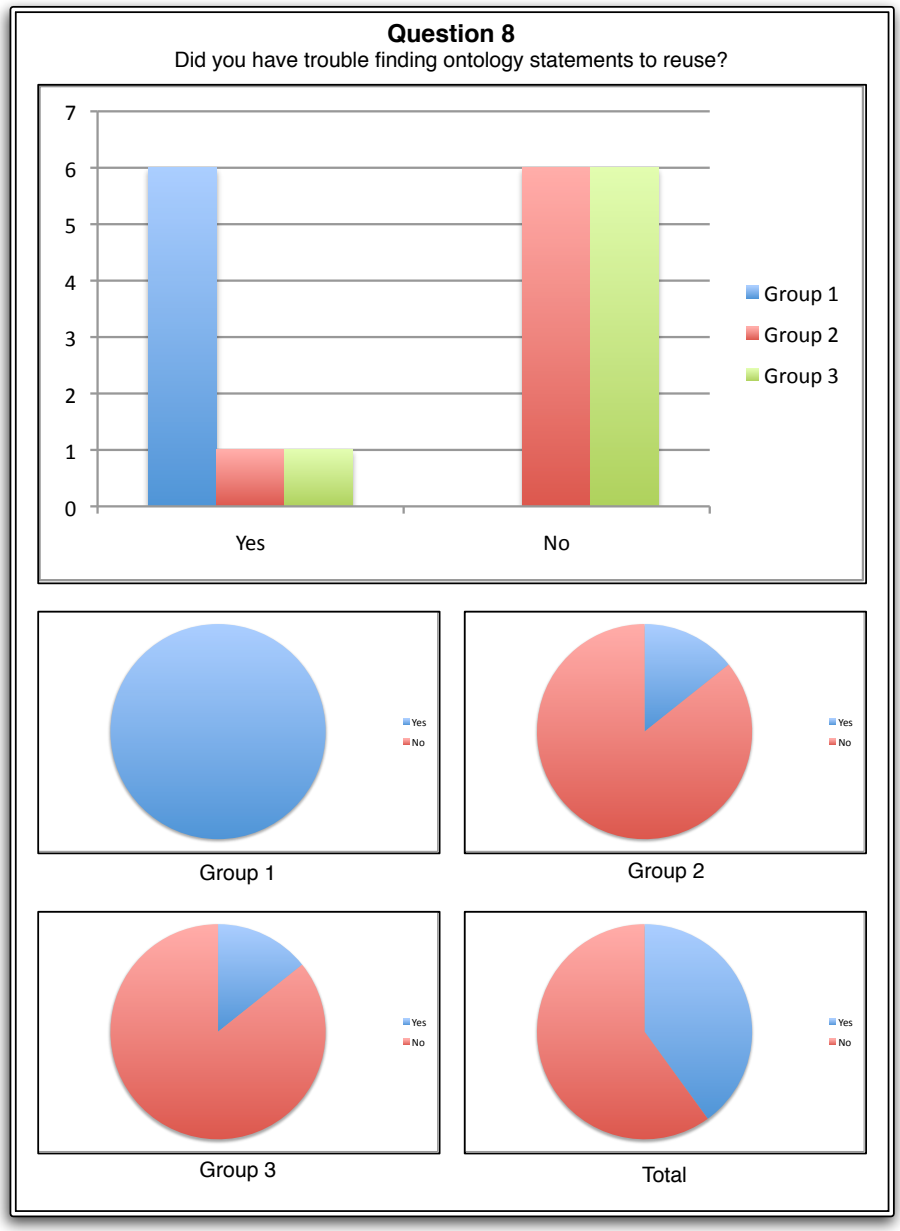

Fig. 11. A graphical representation of the results for question 8.

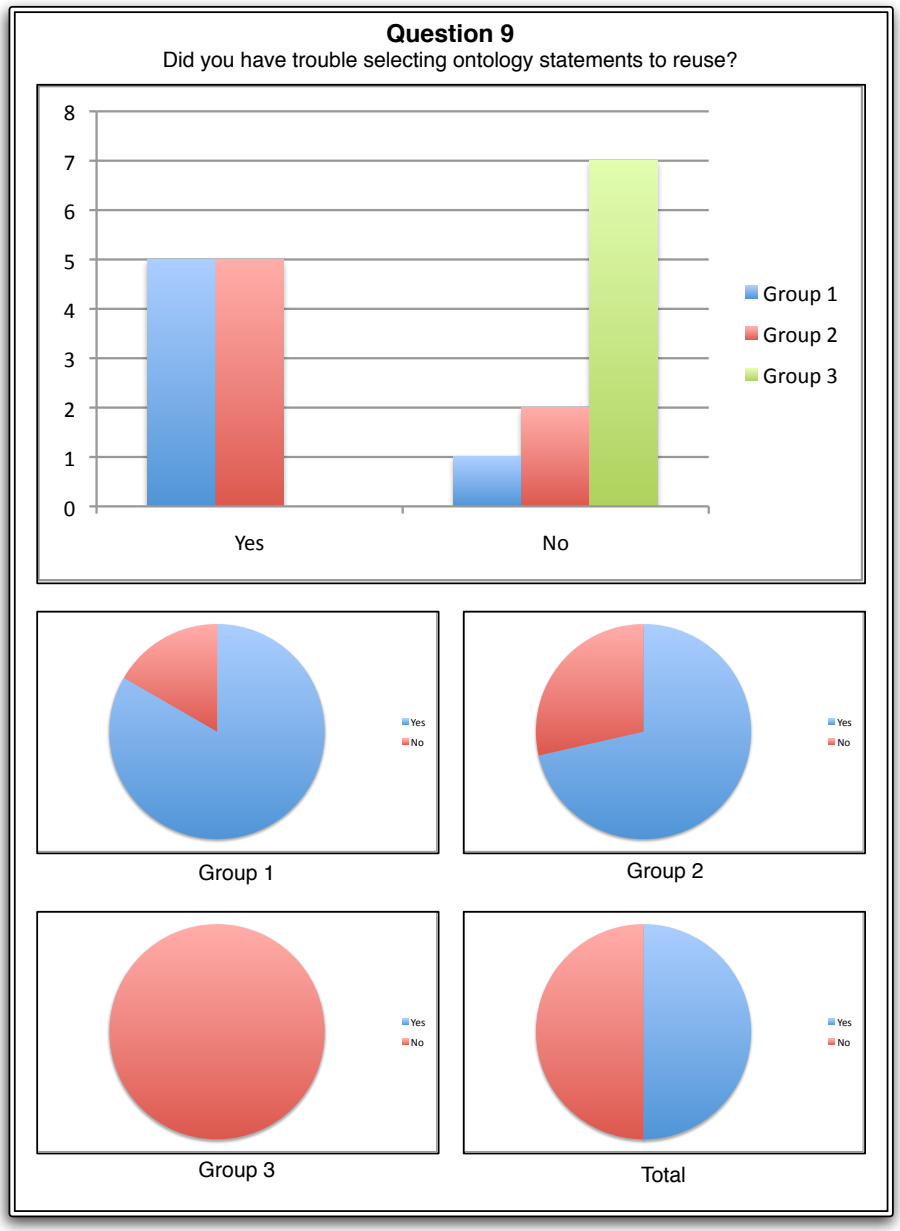

Fig. 12. A graphical representation of the results for question 9.

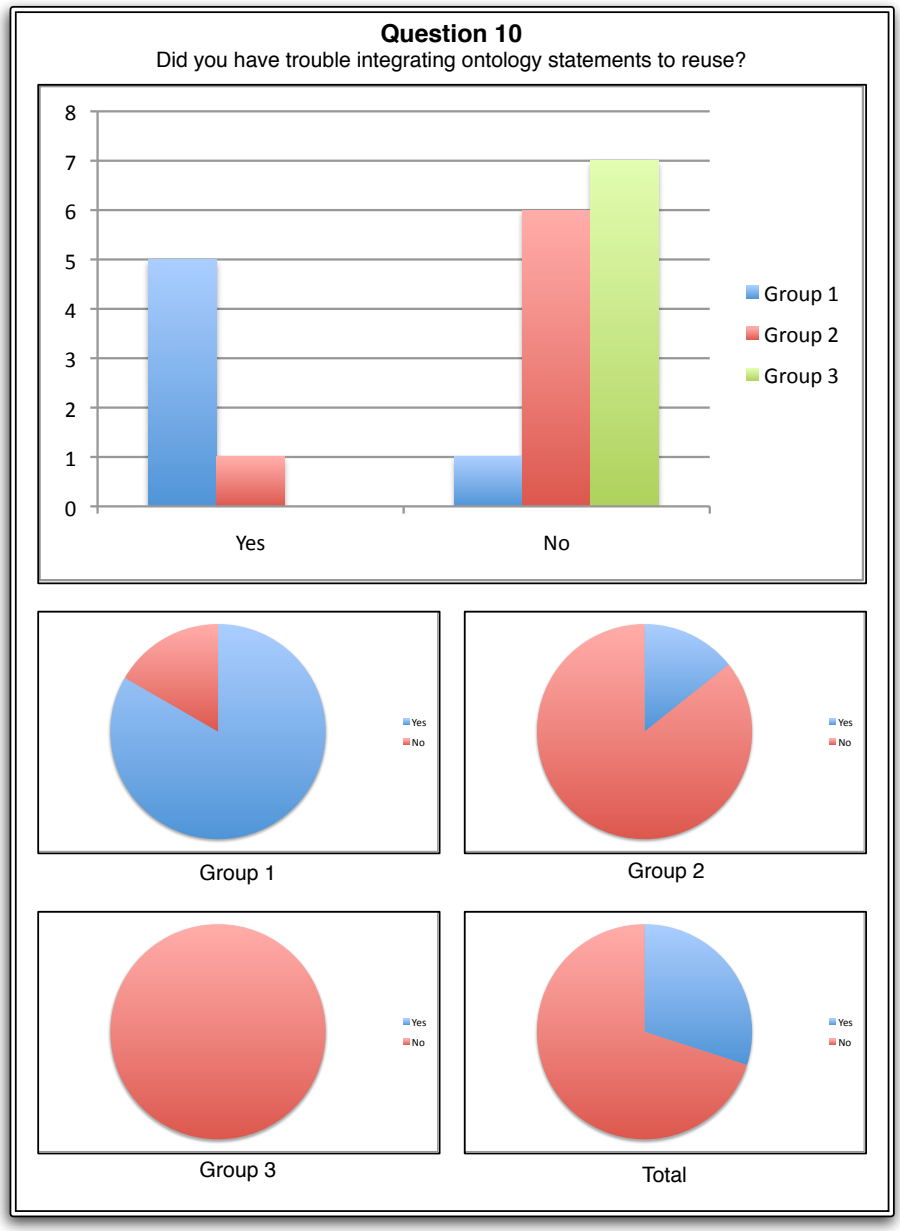

Fig. 13. A graphical representation of the results for question 10.

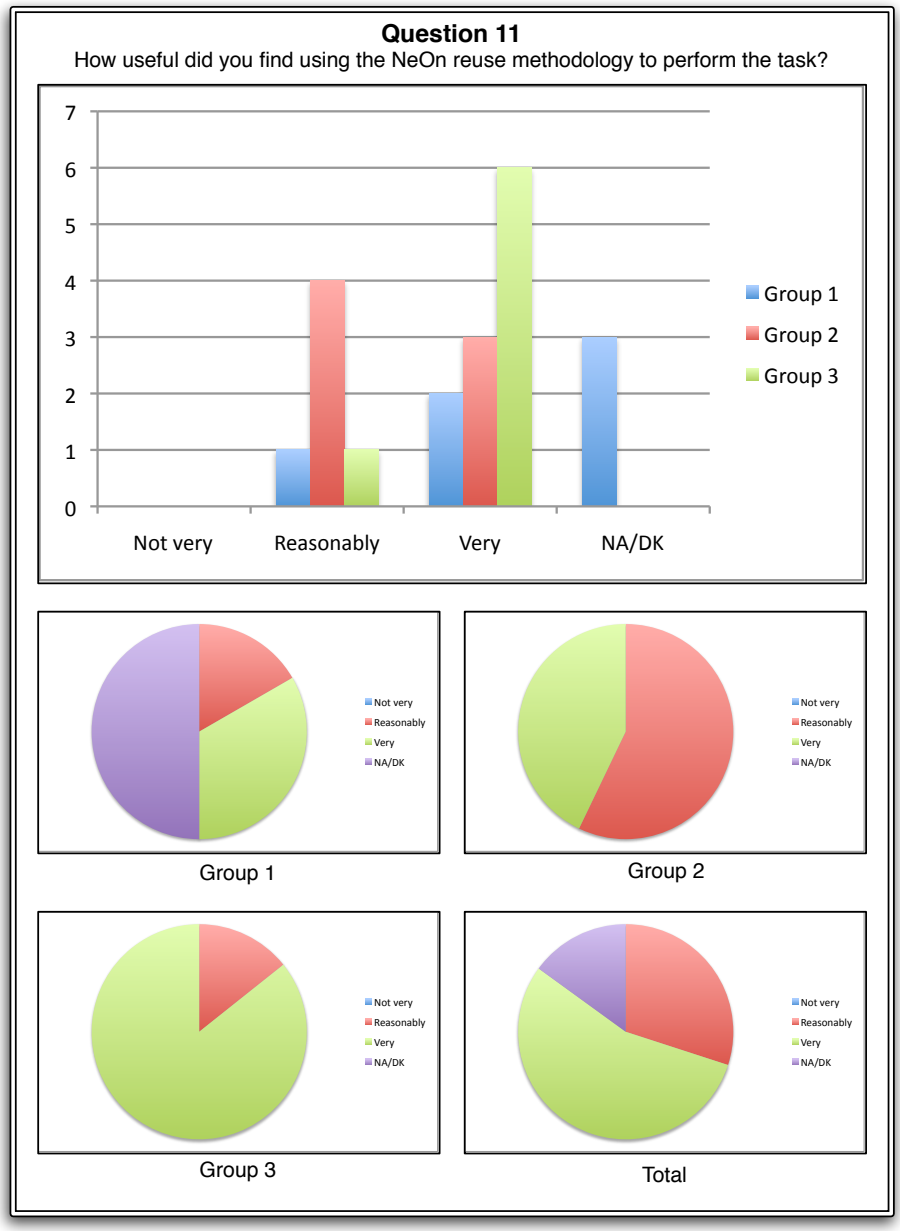

Fig. 14. A graphical representation of the results for question 11.

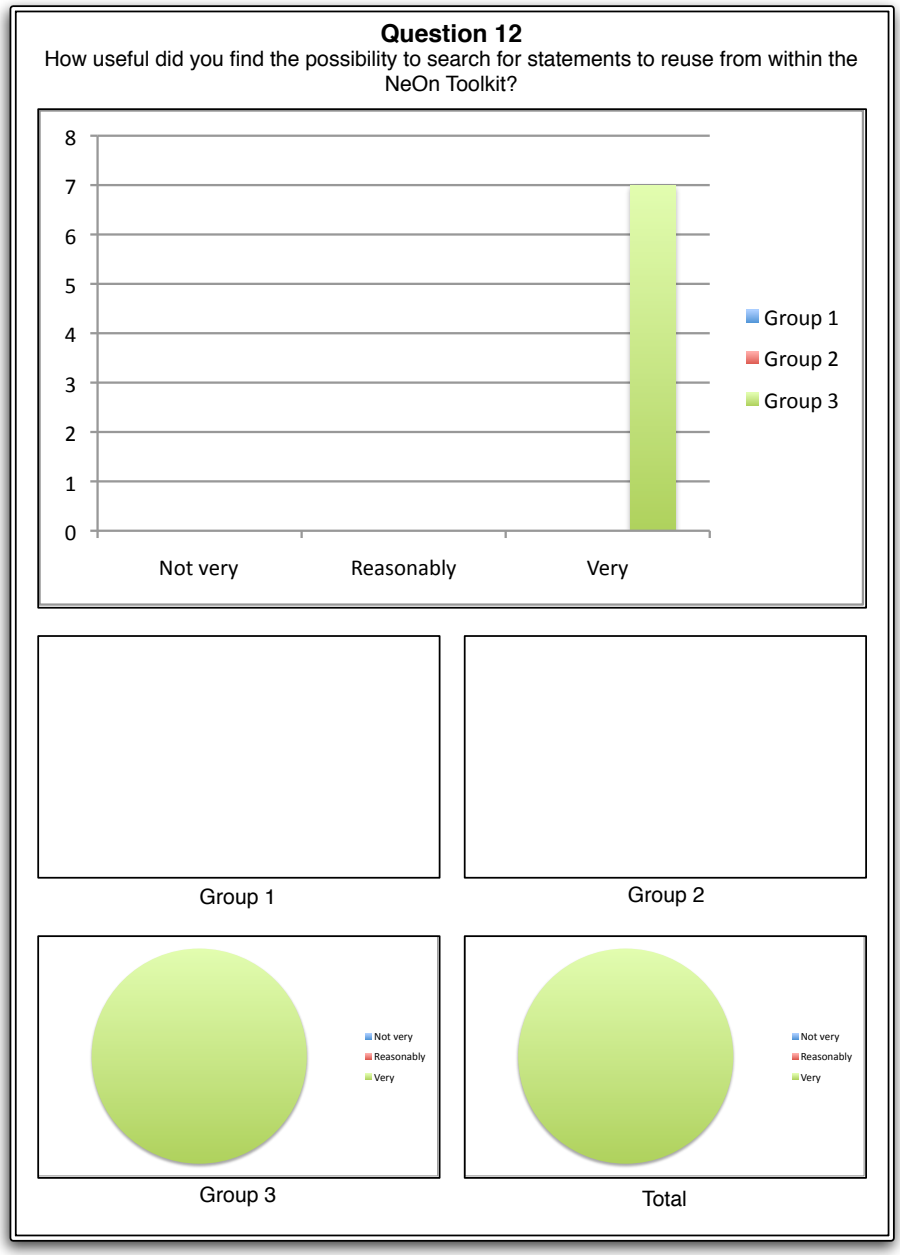

Fig. 15. A graphical representation of the results for question 12.

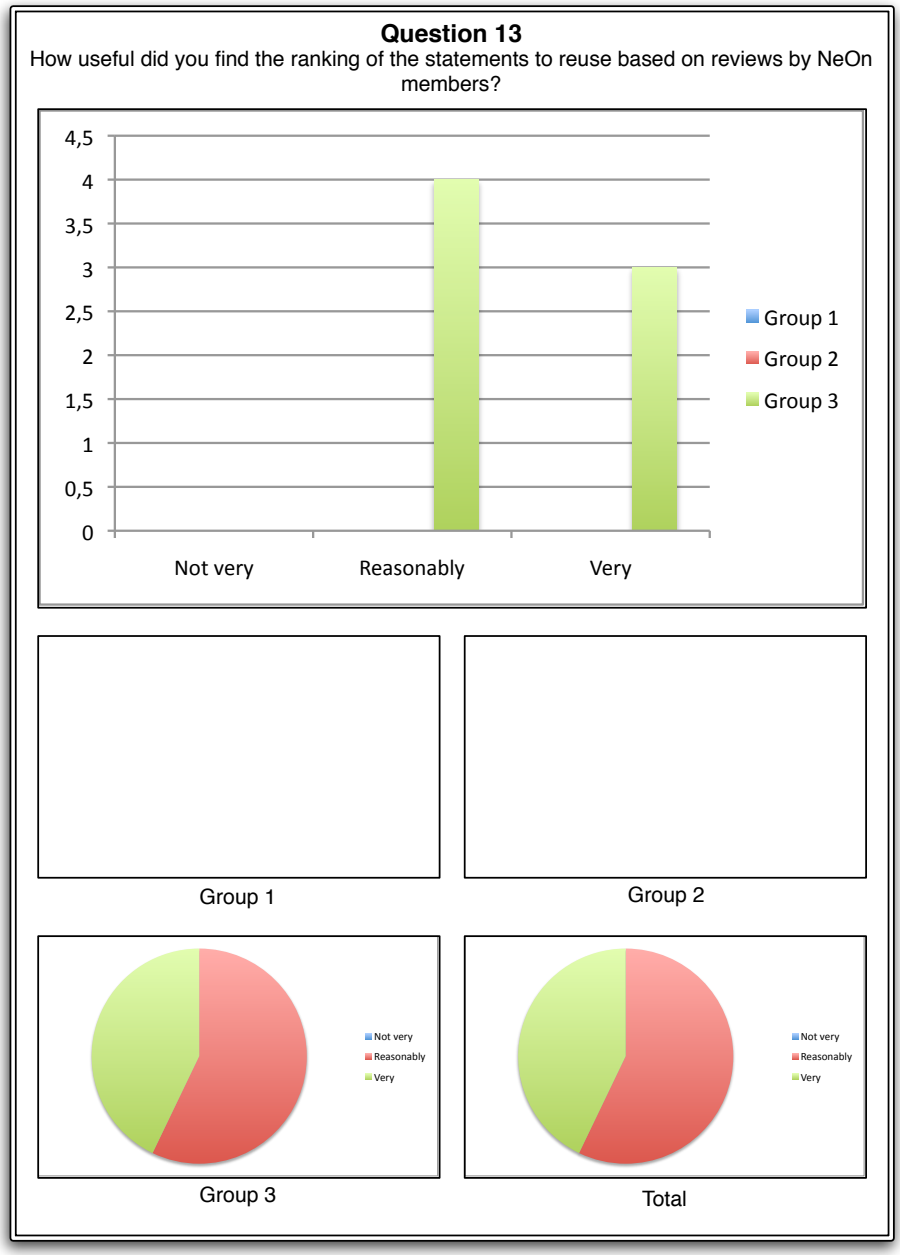

Fig. 16. A graphical representation of the results for question 13.

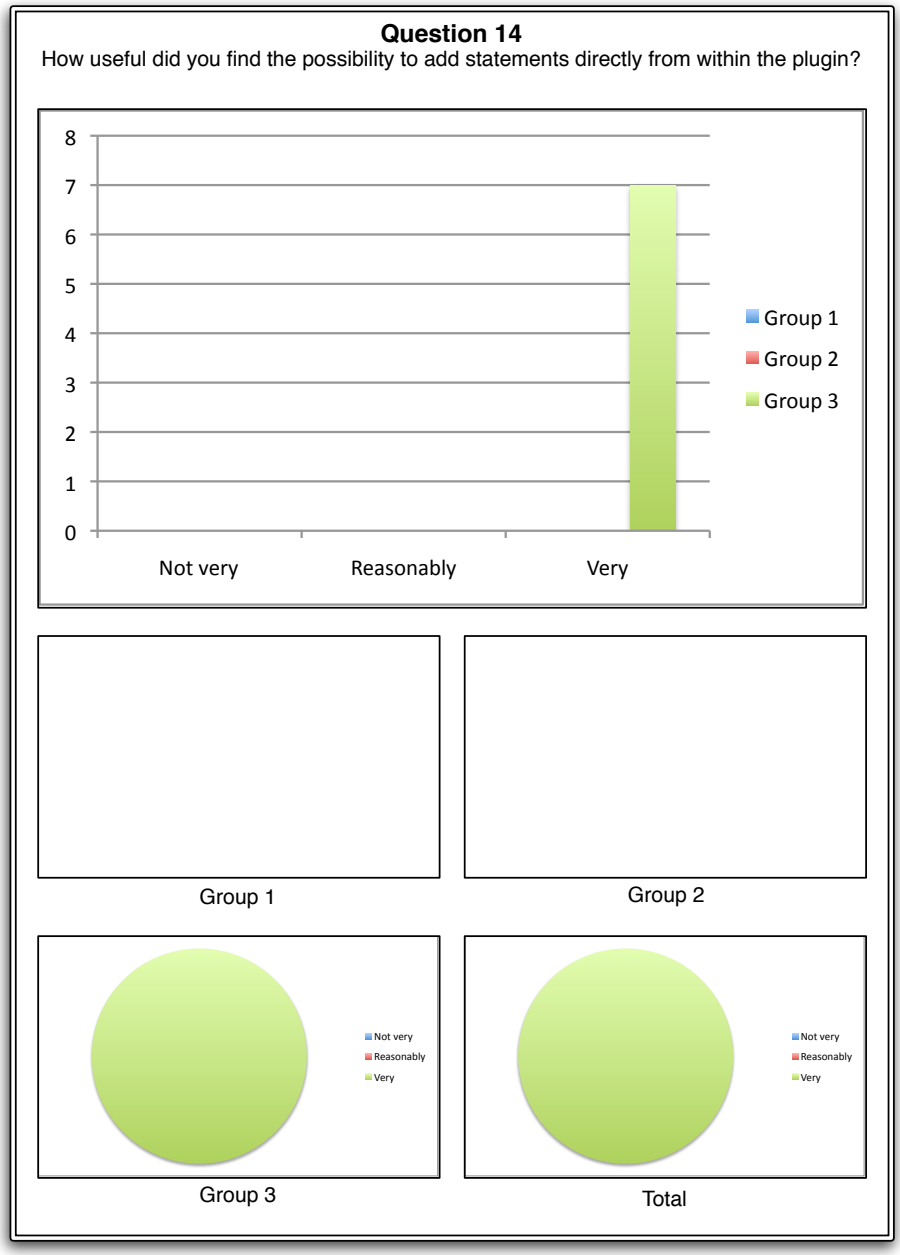

Fig. 17. A graphical representation of the results for question 14.

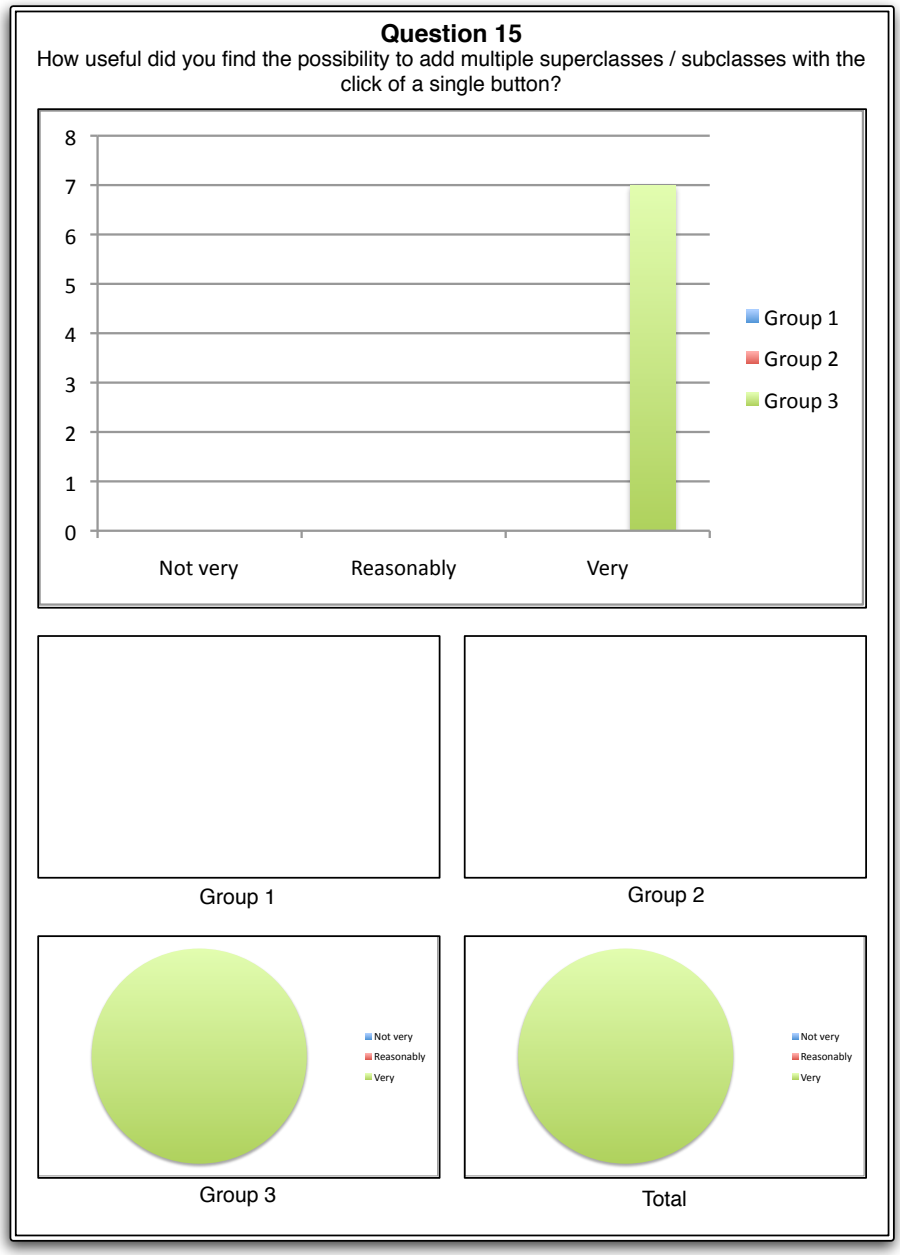

Fig. 18. A graphical representation of the results for question 15.

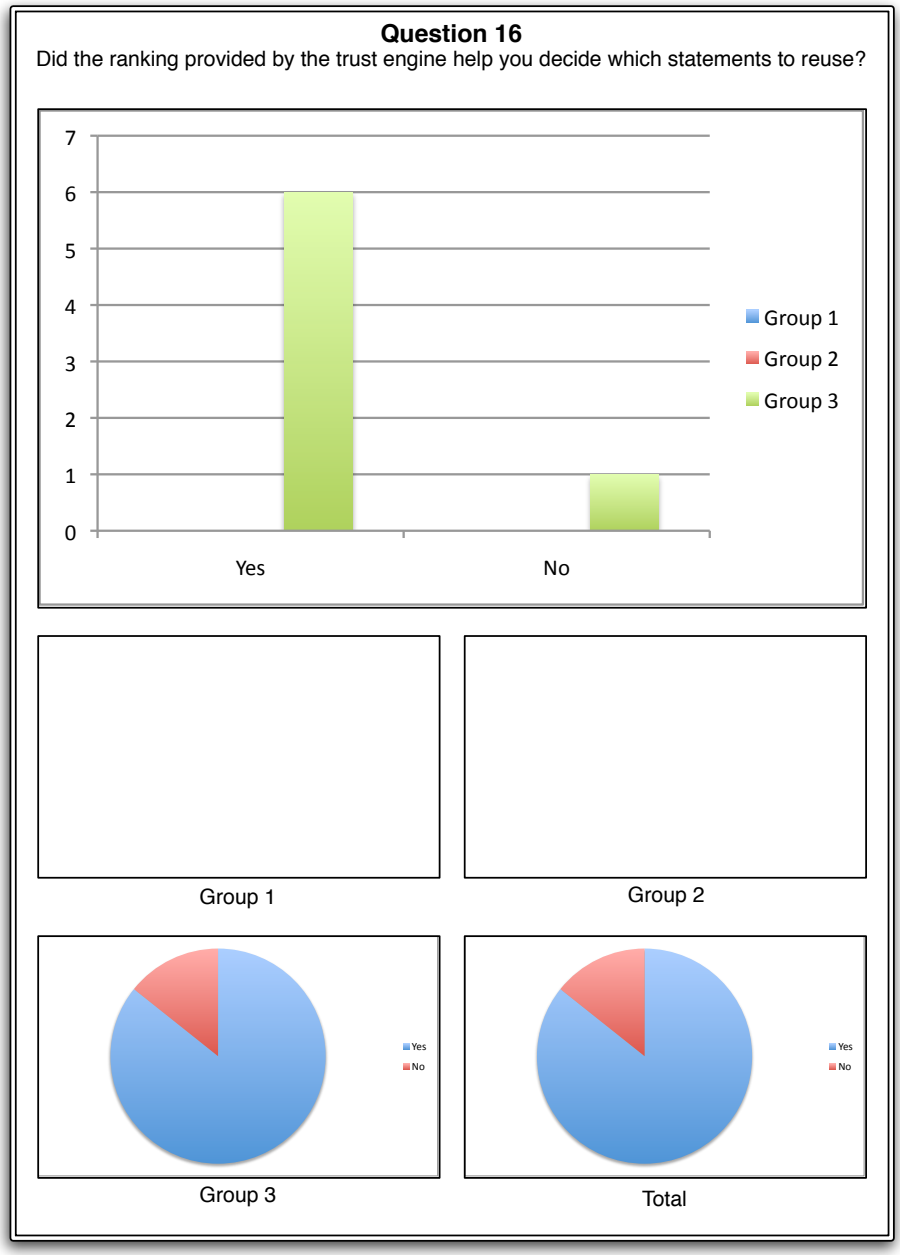

Fig. 19. A graphical representation of the results for question 16.

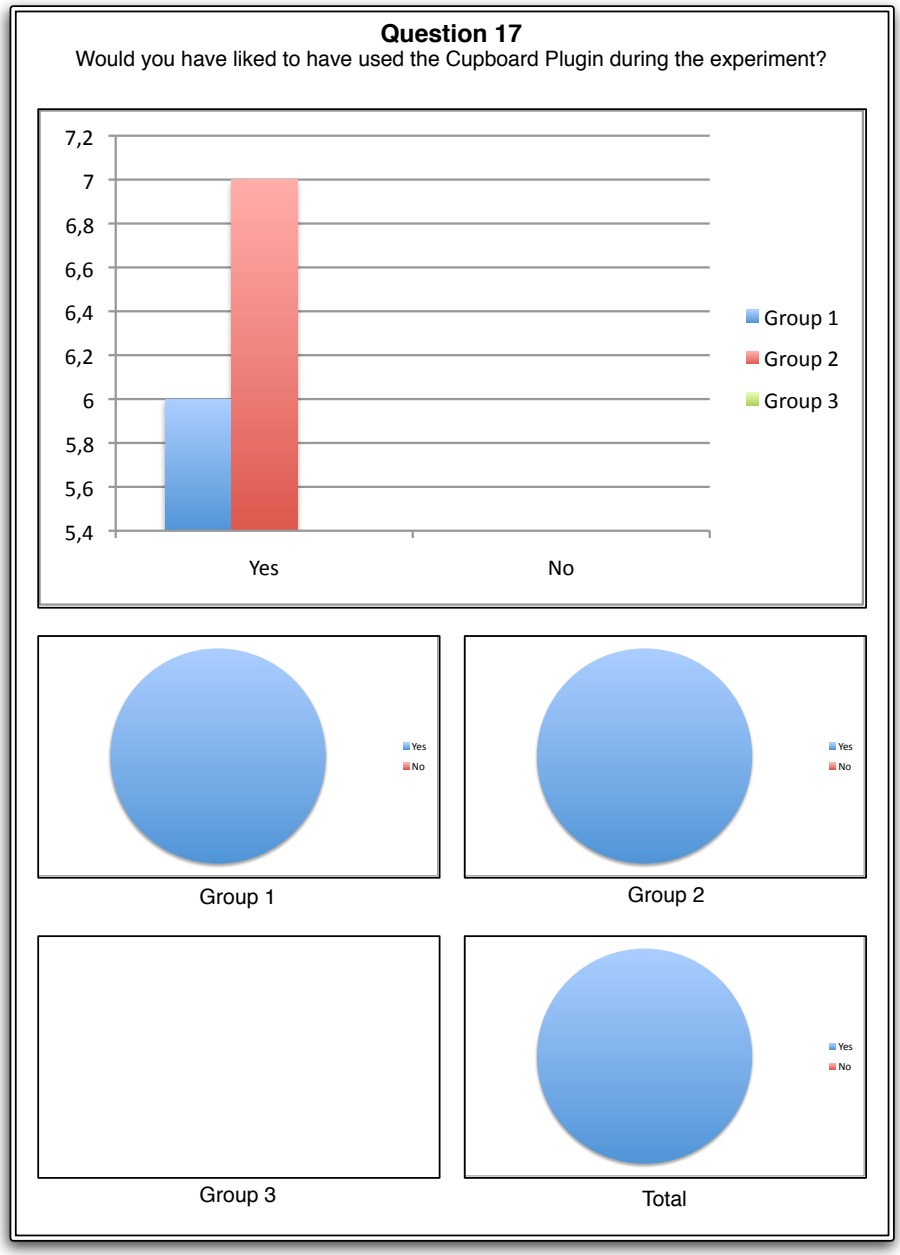

Fig. 20. A graphical representation of the results for question 17.

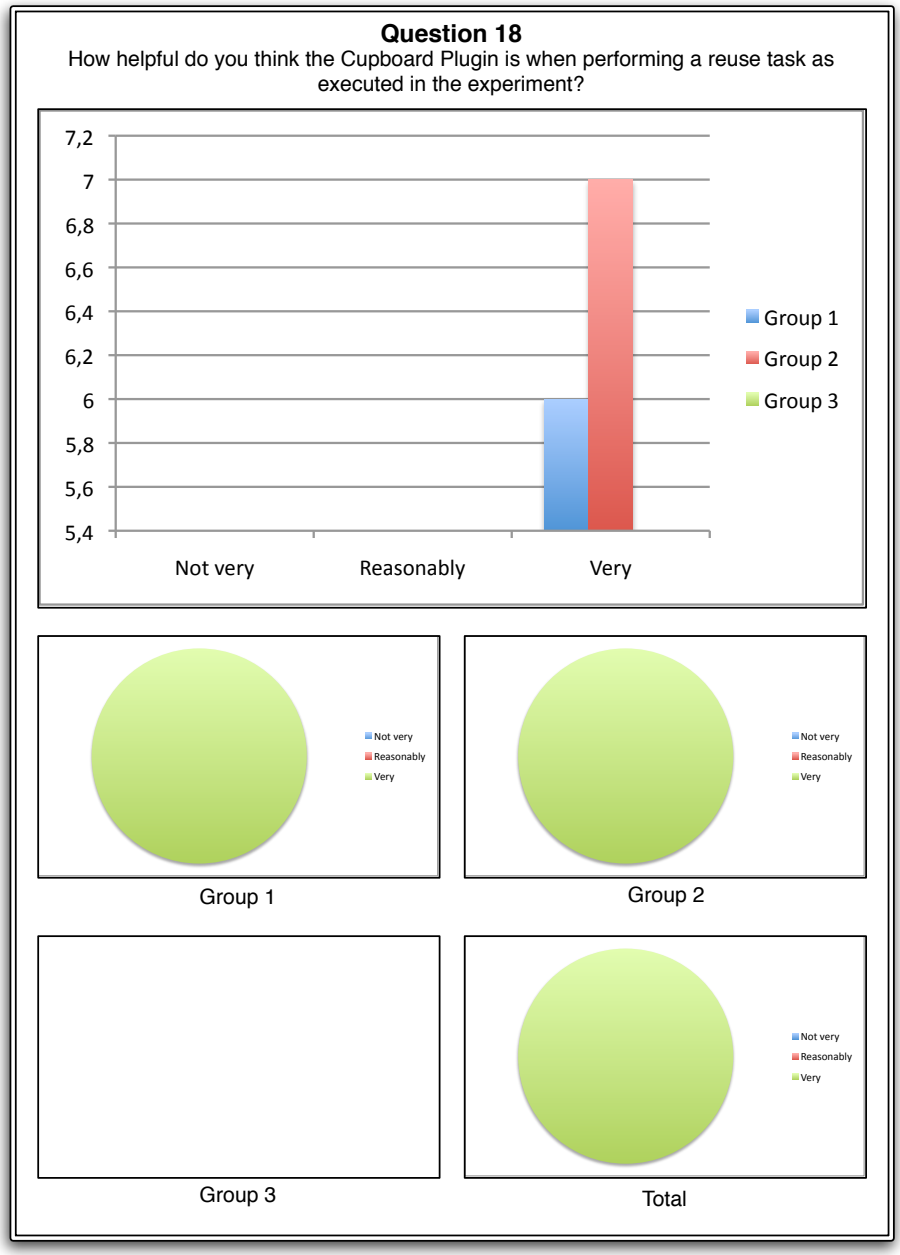

Fig. 21. A graphical representation of the results for question 18.

6.2 Experiment-related Documents

NeOn Ontology Reuse Experiment

Please read all the provided material carefully before beginning the experiment. In case you have questions, please ask the person running the experiment.

Scenario: You want to build an **ontology about fish.** As a starting point, you create a new project and ontology within the NeOn Toolkit, which consists of the class „**Fish**".

Task: Now you should **extend** the ontology by reusing **existing** ontological knowledge. You can add new superclasses or subclasses, new relations, new labels, anything that you like to add from existing resources can be added to the ontology.

Please **use the methodological guidelines** handed to you. They contain hints on where you can search for ontologies or statements on the web, and also explain which tools can assist you in your task. Once you have read everything and have created the project with the ontology containing the class "Fish", you have **20 minutes** to complete.

Goal: Your goal is to come up with an ontology that you consider represents (conceptually models) the **fish domain**, including **superclasses** that help classifying what a fish is (like the SUMO upper level ontology) and **fish-subclasses** (i.e. Salmon subclass of Fish).

After the twenty minutes passed, please **save** the ontology and **complete the questionnaire** we will send to you by email.

Fig. 22. This document was given to all participants of the experiments. It lays out the task they were asked to do.

Ontology Statement Reuse Experiment

Methodological Guidelines

Group 1

Allowed tools: Internet, NeOn Toolkit w/o additional plugins except for RaDON.

Step 1: Ontology Statement Search

In this step, you search the Internet for candidate ontology statements that can be reused in the ontology you want to build (e.g. search in Watson for "fish").

A list of ontology search engines can be found here:

http://tinyurl.com/ontose

Please note that semantic web gateways like Watson allow direct search on the statement level, including information about subclasses and superclasses.

Step 2: Ontology Statement Assessment

In this step you decide which of the ontology statement is useful or not for the ontology being developed or extended. Some criteria to be considered are:

- Does the statement belong to an ontology that covers the same or a similar scope like the ontology being developed
- Check whether the purpose of the statement in the original ontology is similar to the purpose of the ontology developed
- Check the clarity of the ontology statement
- Check the information content of the statement
- Assess the correctness of the statement from a formal modeling perspective

Step 3: Ontology Statement Selection

Select the best statements of the statements found for reuse.

Step 4: Ontology Statement Integration

Integrate the selected statement into the ontology being developed.

Step 5: Check Local Inconsistencies

In the last step the ontology has to be checked for inconsistencies. You can either try to do this manually, or use the RaDON plugin (if unsure how to use it, ask the facilitator).

Fig. 23. The methodological guidelines for group 1.

Ontology Statement Reuse Experiment

Methodological Guidelines

Group 2

Allowed tools: Internet, NeOn Toolkit w/o additional plugins except for RaDON and the Watson plugin.

In order to help you with the process, you can use the Watson plugin to search for statements to reuse and add them to the ontology.

Step 1: Ontology Statement Search

In this step, you search the Internet for candidate ontology statements that can be reused in the ontology you want to build (e.g. search in Watson for "fish").

A list of ontology search engines can be found here:

http://tinyurl.com/ontose

Please note that semantic web gateways like Watson allow direct search on the statement level, including information about subclasses and superclasses.

You can query Watson directly from within the NeOn Toolkit using the Watson plugin.

Step 2: Ontology Statement Assessment

In this step you decide which of the ontology statement is useful or not for the ontology being developed or extended. Some criteria to be considered are:

- Does the statement belong to an ontology that covers the same or a similar scope like the ontology being developed
- Check whether the purpose of the statement in the original ontology is similar to the purpose of the ontology developed
- Check the clarity of the ontology statement
- Check the information content of the statement
- Assess the correctness of the statement from a formal modeling perspective

Step 3: Ontology Statement Selection

Select the best statements of the statements found for reuse.

Step 4: Ontology Statement Integration

Integrate the selected statement into the ontology being developed.

When using the Watson Plugin, statements can be integrated by clicking the add button.

Step 5: Check Local Inconsistencies

In the last step the ontology has to be checked for inconsistencies. You can either try to do this manually, or use the RaDON plugin (if unsure how to use it, ask the facilitator).

Fig. 24. The methodological guidelines for group 2.

Ontology Statement Reuse Experiment

Methodological Guidelines

Group 3

Allowed tools: Internet, NeOn Toolkit w/o additional plugins except for RaDON and the Cupboard plugin

In order to help you with the process, you can use the Cupboard plugin to search for statements to reuse and add them to the ontology. Cupboard will rank the ontologies based on trust and reviews added by members of the NeOn team. You can also use the "add all subclasses" or "add all superclasses" features of the plugin.

Step 1: Ontology Statement Search

In this step, you search the Internet for candidate ontology statements that can be reused in the ontology you want to build (e.g. search in Watson for "fish").

A list of ontology search engines can be found here: http://tinyurl.com/ontose

Using the Cupboard plugin, you can search for statements directly from within the NeOn Toolkit.

Step 2: Ontology Statement Assessment

In this step you decide which of the ontology statement is useful or not for the ontology being developed or extended. Some criteria to be considered are:

- Does the statement belong to an ontology that covers the same or a similar scope like the ontology being developed
- Check whether the purpose of the statement in the original ontology is similar to the purpose of the ontology developed
- Check the clarity of the ontology statement
- Check the information content of the statement
- Assess the correctness of the statement from a formal modeling perspective

Please note that when using the Cupboard plugin, the ontologies come ranked based on reviews from NeOn members. They have reviewed the ontologies fort he task of reusing them in the fish domain. In case you want to see the reviews, you can look here: http://kmi-web06.open.ac.uk:8081/cupboard/Experiment1

Step 3: Ontology Statement Selection

Select the best statements of the statements found for reuse.

When using the Cupboard plugin, the ontology statements are ranked based on the reviews on the ontologies they are contained in. Statements from better ontologies are ranked higher.

Step 4: Ontology Statement Integration

Integrate the selected statement into the ontology being developed.

When using the Cupboad plugin, the statements can be included by simply clicking a button. It is also possible to add all sublasses or all superclasses at the click of one button, to avoid adding all subclasses and searching for them again to find more subclasses.

Step 5: Check Local Inconsistencies

In the last step the ontology has to be checked for inconsistencies. You can either try to do this manually, or use the RaDON plugin (if unsure how to use it, ask the facilitator).

Fig. 25. The methodological guidelines for group 3.

Guidance for Facilitators Running the NeOn Reuse Experiment

First of all, thank you for volunteering to participate in the NeOn Reuse Experiment.

Please make sure that you find **at least three** people in your institution to participate in the experiment and group them in three groups. The different groups will have different tools at their disposal to complete the task given. Timeslot per participant roughly 30-40 minutes.

Group 1:

- Methodological Guidelines for Group 1
- Description of Task to be performed
- A computer with the latest version of the NeOn Toolkit and the RaDON Plugin installed
- Access to the internet
- A sheet of paper to take notes

Group 2:

- Methodological Guidelines for Group 2
- Description of Task to be performed
- A computer with the latest version of the NeOn Toolkit and the RaDON Plugin installed
- The Watson Plugin for the NeOn toolkit installed
- Access to the internet
- A sheet of paper to take notes

Group 3:

- Methodological Guidelines for Group 3
- Description of Task to be performed
- A computer with the latest version of the NeOn Toolkit and the RaDON Plugin installed
- The Cupboard Plugin for the NeOn toolkit installed
- Access to the internet
- A sheet of paper to take notes

Preparation: Depending on the platform you want to run the experiment on, please download the latest version of the NeOn Toolkit and install the RaDON Plugin using the update mechanism. Be sure that for the different groups, users do not have access to the Watson plugin or Cupboard plugin unless specifically mentioned in the instructions above.

Running the experiment:

Fig. 26. First page of the facilitator guidlines.

Please prepare the computer beforehand and ensure that all needed materials are available. Please also make sure recording software is available to capture the screen during the experiments.

Once the participant has been given the material (based on the group you assigned them to), make sure he or she reads it and understands what to do. You are to provide help if needed. Once the document is needed, the ontology project with the ontology containing the class fish has to be created. In case the user does not know how to use the NeOn Toolkit, you should briefly explain the functionality, since usability of the NTK is not tested in the experiment. You can guide the user through the process of creating the initial ontology. For users in group 2 and 3, you should also tell them that the Watson or Cupboard plugin can be invoked using a right-click on the concept und selecting the search functionality from the context menu.

Once the participant says ready, turn on the screen capturing. During the next 20 minutes, the participant should perform the experiment, i.e. searching for reusable content and reusing it. If the user says he is finished, you can stop the experiment before 20 minutes are finished. After 20 minutes, the current state of the ontology should be saved (using as filename "Group-X-INST-User-Y", where X is the user group (1-3), INST is your institution code, like OU or UKARL and Y a incremented number (e.g. 2nd user in this group)) and the screen capturing stopped.

For users of group 1 and group 2, please show them quickly the Cupboard Plugin and which functionality it offers (1-2 minutes).

After that, please send the questionnaire by email to the participant including the filename (Group-X-INST-User-Y) as reference.

After all experiments are conducted, please make the results available to Holger Lewen (hle@aifb.uni-karlsruhe.de). You can also not comments and impressions you had during the experiment.

Thank you very much for your help!

Fig. 27. Second page of the facilitator guidlines.

NeOn Reuse Experiment Questionnaire

1. How would you rate your previous experience with the tools used in the test?

Beginner	*Moderate*	*Expert*	*NA/DK*

2. How would you rate your previous experience in ontology engineering?

Beginner	*Moderate*	*Expert*	*NA/DK*

3. Please indicate how you perceived the amount of time needed to execute the tasks of the experiment:

Low	*Average*	*High*	*NA/DK*

4. Your understanding of the tasks comprised in the experiment was:

Low	*Average*	*High*	*NA/DK*

5. How did you find the support provided by the facilitator?

Inadequate	*Adequate*	*Excellent*	*NA/DK*

6. How would you rate the difficulty of the task you executed?

Low	*Average*	*High*	*NA/DK*

7. Did you use the NeOn Toolkit before?

Yes	*No*

8. Did you have trouble finding ontology statements to reuse?

Yes	*No*

9. Did you have trouble selecting ontology statements to reuse?

Yes	*No*

Fig. 28. First page of the experiment questionnaire.

10. Did you have trouble integrating ontology statements to reuse?

Yes	*No*

11. How useful did you find using the NeOn reuse methodology as a guideline to perform the task?

Not very	*Reasonably*	*Very*	*NA/DK*

■ This part only if you used the Cupboard Plugin during the experiment. If not, go to question 17.

12. How useful did you find the possibility to search for statements to reuse from within the NeOn Toolkit?

Not very	*Reasonably*	*Very*	*NA/DK*

13. How useful did you find the ranking of the statements based on reviews by NeOn members?

Not very	*Reasonably*	*Very*	*NA/DK*

14. How useful did you find the possibility to add statements directly from within the plugin?

Not very	*Reasonably*	*Very*	*NA/DK*

15. How useful did you find the possibility to add multiple superclasses / subclasses with the click of a single button?

Not very	*Reasonably*	*Very*	*NA/DK*

16. Did the ranking provided by the trust engine help you decide which statement to reuse?

Yes	*No*

- Please go to question 19.

■ This part only if the Cupboard Plugin was not used in the experiment, but shown afterwards

17. Would you have liked to have used the Cupboard Plugin during the experiment?

Yes	*No*

18. How helpful do you think the Cupboard Plugin is when performing a reuse task as executed in the experiment?

Not very	*Reasonably*	*Very*	*NA/DK*

Fig. 29. Second page of the experiment questionnaire.

19. What functionalities would you like to see in next versions of the Cupboard Plugin?

20. Please, add any critical comments or positive suggestions on how the system might be improved.

21. Finally, could you add any comments, criticisms or suggestions about any aspect of the system not covered in the above questions? Thanks for your cooperation in this.

Fig. 30. Third page of the experiment questionnaire.

References

1. d'Aquin, M., Motta, E., Dzbor, M., Gridinoc, L., Heath, T., Sabou, M.: Collaborative semantic authoring. Intelligent Systems, IEEE **23**(3) (May-June 2008) 80–83
2. d'Aquin, M., Lewen, H.: Cupboard - A Place to Expose Your Ontologies to Applications and the Community. In Aroyo, L., Traverso, P., Ciravegna, F., Cimiano, P., Heath, T., Hyvönen, E., Mizoguchi, R., Oren, E., Sabou, M., Simperl, E.P.B., eds.: ESWC. Volume 5554 of Lecture Notes in Computer Science., Springer (2009) 913–918
3. Lewen, H., Supekar, K., Noy, N.F., Musen, M.A.: Topic-Specific Trust and Open Rating Systems: An Approach for Ontology Evaluation. In: Proc. of the 4th International Workshop on Evaluation of Ontologies for the Web (EON2006) at the 15th International World Wide Web Conference (WWW). (MAY 2006)
4. Sabou, M., Angeletou, S., d'Aquin, M., Barrasa, J., Dellschaft, K., Gangemi, A., Lehmann, J., Lewen, H., Maynard, D., Mladenic, D., Nissim, M., Peters, W., Presutti, V., Villazon, B.: D2.2.1 methods for selection and integration of reusable components from formal or informal user specifications. NeOn Project Deliverable D2.2.1, The Open University (MAY 2007)
5. Gangemi, A., Catenacci, C., Ciaramita, M., Lehmann, J.: Modelling ontology evaluation and validation. In Sure, Y., Domingue, J., eds.: ESWC. Volume 4011 of Lecture Notes in Computer Science., Springer (2006) 140–154