2019

On the Mental Workload Assessment of Uplift Mapping Representations in Linked Data

Ademar Crotti Junior

Christophe Debruyne

Luca Longo

See next page for additional authors

Funder: CNPQ, National Counsel of Technological and Scientific Development – Brazil; Science Foundation Ireland

Authors
Ademar Crotti Junior, Christophe Debruyne, Luca Longo, and Declan O'Sullivan

On the Mental Workload Assessment of Uplift Mapping Representations in Linked Data

Conference Paper · February 2019

DOI: 10.1007/978-3-030-14273-5_10

4 authors:

Ademar Crotti Junior
Trinity College Dublin

Christophe Debruyne
University of Liège

Luca Longo
Technological University Dublin - City Campus

Declan O'Sullivan
Trinity College Dublin

On the Mental Workload Assessment of Uplift Mapping Representations in Linked Data

Ademar Crotti Junior[1][0000-0003-1025-9262], Christophe Debruyne[1][0000-0003-4734-3847], Luca Longo[2][0000-0002-2718-5426], Declan O'Sullivan[1][0000-0003-1090-3548]

[1] ADAPT Centre, Trinity College Dublin, Dublin 2, Ireland
{ademar.crotti, christophe.debruyne,declan.osullivan}@adaptcentre.ie
[2] ADAPT Centre, Dublin Institute of Technology, Dublin 8, Ireland
luca.longo@dit.ie

Abstract. Self-reporting procedures have been largely employed in literature to measure the mental workload experienced by users when executing a specific task. This research proposes the adoption of these mental workload assessment techniques to the task of creating uplift mappings in Linked Data. A user study has been performed to compare the mental workload of "manually" creating such mappings, using a formal mapping language and a text editor, to the use of a visual representation, based on the block metaphor, that generate these mappings. Two subjective mental workload instruments, namely the NASA Task Load Index and the Workload Profile, were applied in this study. Preliminary results show the reliability of these instruments in measuring the perceived mental workload for the task of creating uplift mappings. Results also indicate that participants using the visual representation achieved smaller and more consistent scores of mental workload.

Keywords. Mental Workload; Uplift Mapping Representations; Linked Data.

1 Introduction

Human mental workload (MWL) is a fundamental design concept used to investigate the interaction of human with computers and other technological devices [22]. MWL instruments measure the cognitive load experienced by users when executing a specific task [5]. Literature suggests that both mental overload and underload can affect performance [22].

This study employs human mental workload instruments to the task of creating uplift mappings in Linked Data. Linked Data refers to a set of best practices for publishing and interlinking data on the Web [4]. The standard data model used in Linked Data is the Resource Description Framework[1] (RDF). Uplift mappings are responsible for expressing how non-RDF data should be transformed to RDF [8]. A significant part of the Linked Data web is achieved by such conversion process.

The uplift process is often express through mapping languages. The W3C Recommendation mapping language R2RML [9] (RDB to RDF mapping language) is

[1] http://www.w3.org/TR/rdf11-concepts/

an example of a formal language used to express mappings that transform relational databases into RDF. These mappings can be created “manually”, trough text editors or by applications that support user involvement in the mapping process. Such applications may make use of visual representations to alleviate the knowledge required by mapping languages [31]. An example of a visual representation is the **J**igsaw P**u**zzles for Representing **Ma**ppings (Juma) [18]. Juma is based on the block metaphor, which has become popular with visual programming languages (see 2.3).

It is assumed that the creation of mappings using different uplift mapping representations require different cognitive processing resources. And that the assessment of the cognitive workload of uplift mapping representations can be used to evaluate and improve the interaction between users and these representations.

Thus, this paper extends the application of MWL instruments by evaluating the perceived mental workload of users when performing an uplift mapping task. The user experiment presented in this paper assesses the cognitive load of creating uplift mappings using the two aforementioned mapping representations, R2RML and Juma. Two subjective mental workload instruments were applied in this study, namely the Workload Profile and the Nasa Task Load Index. To the authors knowledge, this paper presents the first evaluations considering the cognitive load of creating uplift mappings in Linked Data.

The remainder of this paper is structured as follows: Section 2 discusses the background knowledge, which contains a brief description of mappings applied in the Linked Data domain. Section 3 presents the two mental workload assessment instruments used in this study. Section 4 introduces the design of a novel primary research at the intersection of mental workload and uplifting mapping tasks. Results and their analysis are presented in Section 5. Related work is presented in Section 6. Section 7 concludes the paper and suggests future work.

2 Background

2.1 Mappings in Linked Data

The term Linked Data refers to a set of best practices for publishing and interlinking data on the Web [4]. A Linked Data dataset is structured information encoded using the Resource Description Framework (RDF), that are linked to other datasets, and accessible via HTTP. RDF is a graph data model that provides one means to describe, annotate and exchange information such that machines can process them [4].

The Linking Open Data project has the goal of publishing open datasets as Linked Data. These open datasets are freely accessible and collectively known as the Linked Open Data cloud[2]. A significant part of the Linked Data cloud is achieved by converting resources to RDF, often through mappings. In a general context, a mapping defines a relation between source and target elements [12]. The properties of a mapping are represented in a structured format using mapping languages [8]. Mappings that express how non-RDF data is transformed to RDF are called uplift mappings. An example of a transformation from a relational database to RDF is presented in **Fig. 1**. In this example, the table *person* is transformed into the graph-based RDF data model.

[2] http://lod-cloud.net/

Fig. 1. Example of a transformation from a relational database to RDF

The R2RML mapping language, which can be used to express these transformations, is presented in Section 2.2. Juma, a visual representation that can be used to generate such mappings, is presented in Section 2.3.

2.2 R2RML

The RDB to RDF mapping language (R2RML) [9] is the W3C Recommendation mapping language used to express mappings between relational databases and RDF. R2RML's vocabulary defines that each mapping consists of one or more *triples maps*. A triples map has (1) one *logical table*, (2) one *subject map* and (3) zero or more *predicate object maps*, where:

1. **Logical Table**: a table or an SQL query from which RDF will be generated.
2. **Subject Map**: subject maps define the subjects of the RDF triples. These subjects can be IRIs or blank nodes. One also may specify zero or more URI class types.
3. **Predicate Object Map**: each predicate object map defines the predicates, using predicate maps, and objects, using object maps, of the RDF triples. Each predicate object map must have at least one predicate map and one object map. Predicates must be valid IRIs. Objects can be IRI's, blank nodes or literal values. For literal values, it is possible to define a data type or a language. One may link triples maps using *parent triples map*. A parent triples map can have zero or more join conditions.

Listing 1 shows an example the transformation presented in **Fig. 1** expressed using the R2RML mapping language.

```
<#TripleMap1>
  rr:logicalTable [
   rr:tableName "person";
  ];

  rr:subjectMap [
    rr:template "http://example.org/person/{id}";
    rr:class foaf:Person;
  ];

  rr:predicateObjectMap [
    rr:predicateMap [ rr:constant foaf:name; ];
    rr:objectMap [ rr:column "name"; ];
  ];.
```

Listing 1. R2RML mapping definition

In this mapping, the logical table is defined as *person.* Using one triples map, we define the subjects to have the following URI `http://example.org/person/{id}`. *Id* is an attribute coming from the table *person*. In this sense, for row with *id* equals to 1, this mapping would generate triples with the subject as `http://example.org/person/1`, and so on. A class definition construct is used to define that these subjects are instances of the class `foaf:Person`, which is declared in the FOAF[3] vocabulary. A predicate object map defines the predicate of the triples to be `foaf:name`, and the object of the triples to be come from the attribute *"name"* of the declared logical table *person*. The output of this mapping, considering that the fictional table *person* has only one record with the attribute *id* as an integer with value *1* and attribute *name* as a string with value *"Ana"*, is shown in RDF Turtle syntax in **Listing 2**.

```
<http://example.org/person/1>
        a       <http://xmlns.com/foaf/0.1/Person> ;
        <http://xmlns.com/foaf/0.1/name>
                "Ana" .
```

Listing 2. RDF output from executing the mapping presented in Listing 1

2.3 Juma

Juma is a method for visually representing mappings in Linked Data. Juma is based on the block (or jigsaw) metaphor that has become popular with visual programming languages – where it is called the block paradigm – such as Scratch[4]. This metaphor allows users to focus on the logic instead of the language's syntax. In addition, the block metaphor has been successfully used in other domains [3, 6]. The implementation of Juma applied to uplift languages used in this study is called Juma Uplift [19]. In Juma Uplift, each mapping defines an input source that is associated to 0 or more vocabularies. These vocabularies are then used in the mapping definitions. A mapping is also associated with 0 or more subject definitions. These subject definitions express how subjects are generated from the input data. Each subject definition has associated predicate object definitions. Subject definitions can also declare these to be instances of 0 or more classes, to be a blank node, and associate triples to a named graph. For more information about Juma Uplift the reader is referred to [19]. **Fig. 2** shows the mapping from **Listing 1** represented using the Juma Uplift representation. The RDF output of this mapping was presented in **Listing 2**.

[3] `http://xmlns.com/foaf/0.1/`
[4] `https://scratch.mit.edu/`, last accessed May 2018

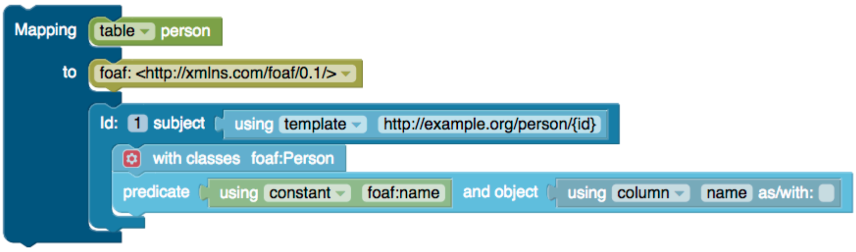

Fig. 2. Juma Uplift mapping representation

3 Mental workload self-reporting assessment instruments

Human mental workload (MWL) is a fundamental design concept used to investigate the interaction of human with computers and other technological devices [22]. It can be intuitively described as the amount of work necessary for users to complete a task [5]. MWL measurements can be classified into three broad categories:

- subjective measures: subjects auto-assess their mental workload by rating a set of dimensions, within pre-defined scales, in relation with the execution of a task performed immediately before;
- performance measures: subjects have some physiological characteristics measured while performing a task. As, for instance, eye activity and heart rate;
- physiological measures: subjects' mental workload is assessed according with the performance reached in a primary or for a secondary task (e.g. error rates; task completion time).

This paper focuses on two subjective mental workload assessment techniques: the Workload Profile and the NASA Task Load Index.

3.1 Workload Profile

The Workload Profile (WP) assessment procedure [43] is built upon the Multiple Resource Theory proposed in [47, 46]. In this theory, individuals are seen as having different capacities or 'resources' related to:

- stage of information processing: perceptual/central processing and response selection/execution;
- code of information processing: spatial/verbal;
- input: visual and auditory processing;
- output: manual and speech output.

Each dimension is quantified through subjective rates and subjects, after task completion, are required to rate the proportion of attentional resources used for performing a given task with a value in the range $0..1 \in \Re$. A rating of 0 means that the task placed no demand while 1 indicates that it required maximum attention. The questionnaire is presented in Table 7. The aggregation strategy is a simple sum of the 8 rates d (averaged here, and scaled in $[1..100 \in \Re]$ for comparison purposes):

$$WP: [0..100] \in \Re$$

$$WP = \frac{1}{8}\sum_{i=1}^{8} d_i * 100$$

3.2 NASA Task Load Index

The NASA Task Load Index (NASA-TLX) instrument [16] belongs to the category of self-assessment measures. It has been validated in the aviation industry and other contexts in Ergonomics [16, 36] with several applications in many socio-technical domains. It is a combination of six factors believed to influence MWL (full questionnaire in Table 8). Each factor is quantified with a subjective judgement coupled with a weight computed via a paired comparison procedure. Subjects are required to decide, for each possible pair (binomial coefficient, $\binom{6}{2} = 15$) of the 6 factors, 'which of the two contributed the most to mental workload during the task', such as 'Mental or Temporal Demand?', and so forth. The weights *w* are the number of times each dimension was selected. In this case, the range is from 0 (not relevant) to 5 (more important than any other attribute). The final MWL score is computed as a weighted average, considering the subjective rating of each attribute d_i and the correspondent weights w_i:

$$NASATLX: [0..100] \in \Re$$
$$NASATLX = \left(\sum_{i=1}^{6} d_i * w_i\right)\frac{1}{15}$$

Alternatively, it is possible to calculate the MWL scores eliminating the weighted procedure, which is called *Raw TLX*.

4 Design and Methodology

A primary research study has been designed to assess the mental workload of creating uplift mappings in Linked Data using two different mapping representations. This experiment compares the "manual" creation of uplift mappings with R2RML using the RDF TURTLE notation[5] (which is in essence a text file) to the visual mapping representation Juma. For the remainder of this paper, R2RML mappings refers to mappings in R2RML using RDF TURTLE syntax, and Juma refers to mappings represented using the Juma Uplift representation.

The research hypotheses related to this experiment are:

- **Hypothesis H1**: the perceived mental workload of users interacting with Juma for the creation of uplift mappings is expected to be lower than the perceived

[5] TURTLE is only one of the many standardized RDF representations. TURTLE was chosen as it is terse, and one of the more usable and easier to read representations. Even the R2RML W3C Recommendation uses TURTLE for their examples.

mental workload experienced by users that crafted the same mappings manually, according to the NASA-TLX and WP mental workload measures.
- **Hypothesis H2**: the NASA-TLX and WP mental workload measures have high reliability.

4.1 Participants and procedure

A number of students enrolled in a third-level class from a MSc module in Information and Knowledge Architecture in Trinity College Dublin, Ireland, in 2017, have been approached for this experiment. The experiment was executed in week 10 of a 12-week module. At that time, the course on Knowledge Engineering and Semantic Web technologies had covered OWL modeling, RDF, and SPARQL (amongst others). Participants also had one class, a week before the experiment, on R2RML, which included exercises. This highlights the pre-training on R2RML that the participants have received prior to this research experiment. Note that participants had no knowledge of Juma prior to the experiment.

In order to evaluate the Juma and R2RML mapping representations for the task of creating uplift mappings, participants were split into two groups. Students in one group were exposed to the Juma visual representation – which, for the remainder of the paper, we refer to as the Juma group. Participants in the second group were able to use their preferred text editor to create uplift mappings manually, using R2RML – referred as the R2RML group for the remainder of the paper.

The study was executed with 26 participants, 12 in the Juma group and 14 in R2RML group. The experiment was executed with participants in a classroom; and lasted for 50 minutes. The first 10 minutes were used to explain the experiment to participants, and for participants to examine the material provided. Note that participants still did not have access to the uplift mapping task at this point. Participants were also asked to fill in, read, and consent to the study information sheet, to be able to participate in the experiment. All participants had exactly 30 minutes for the execution of the task. Finally, in the last 10 minutes, participants were asked to fill in the questionnaires associated to the WP and NASA-TLX mental workload assessment instruments. Note that the question of the NASA−TLX related to 'physical demand' (NT_2 in Table 8) was set to 0, as there is no physical load related to the task assessed in this experiment. In detail, the evaluation was structured in four parts, as also depicted in **Fig. 3**:

1. **Technical debriefing:** all participants had the opportunity to watch videos about R2RML[6] prior to executing the uplift mapping task. The group using the Juma method also had a presentation and a video about the visual representation[7]. The material was also available during the execution of the task.
2. **Mapping task:** in the main part of this study, participants were asked to create a specific uplift mapping (described in section 4.2). Participants could ask questions for clarifying any doubts about the experiment.

[6] Available at `https://www.scss.tcd.ie/~crottija/juma/r2rml.pdf` and `https://www.youtube.com/watch?v=fn5mKGGj2us`.
[7] Available at `https://www.scss.tcd.ie/~crottija/juma/juma.pdf` and `https://www.youtube.com/watch?v=Q97YeZtu_tA`.

3. **Post-task questionnaire:** after completion of the task, participants were asked to fill in the WP and NASA-TLX mental workload questionnaires.

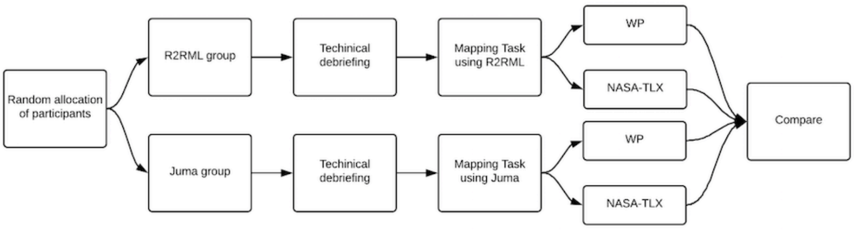

Fig. 3. Experiment design diagram

4.2 Mapping Task

This user study was built on top of the Microsoft Access 2010 Northwind sample database that has been ported to MySQL[8]. Participants were asked to create one R2RML mapping divided in three subtasks. For each subtask, a sample RDF output was shown to participants. In addition, they could run the mapping, by using an R2RML processor, and compare the output of their mapping execution to the sample provided. In this sense, an R2RML processor [10] was integrated to Juma. Participants creating the mappings using a text editor had access to a compacted folder with the same engine and the command line instruction that runs it. By executing the mappings, participants were able to validate the correctness of the output. A summary of the mapping task, separated into its subtasks, is shown below:

- **Subtask 1**: participants had to define a mapping with one subject per row of the table *employees*. The subject URI for the triples should be `http://data.example.org/employee/{id}`. These subject should also have the URI type class `foaf:Person` from the FOAF[9] vocabulary. The mapping definition should also create, for these subjects, the predicate `foaf:givenName` with object from the column *first_name*. The predicate `foaf:familyName` with object from the column *last_name*. Finally, the predicate `foaf:name` should have the concatenation of the columns *last_name* and *first*_name separated by comma as object.
- **Subtask 2**: in the same mapping, participants were asked to define another subject from the table *employees*. The subject URI should be `http://data.example.org/city/{city}`. These subjects should have the URI type class `foaf:Spatial_Thing`. The mapping should generate the predicate `rdfs:label`, from the RDFS[10] vocabulary, with object from the column *city* for each subject.

[8] Available at `https://github.com/dalers/mywind`
[9] `http://xmlns.com/foaf/0.1/`
[10] `http://www.w3.org/2000/01/rdf-schema#`

- **Subtask 3**: finally, participants were asked to link the subject from subtask 1 with the subject from subtask 2 using the predicate `foaf:based_near`.

Some elements of the task could be achieved in different ways. For example, since not all attributes are mapped, participants could map an SQL query instead of the whole table. Concatenating could be implemented using a template construct, an SQL query, or through the use of the data transformation function called ‘concatenating’ - for participants using Juma Uplift. The template construct would be the expected solution to concatenating. Subtask 3 asked participants to relate the subjects created in subtask 1 and subtask 2. This could be achieved by mapping using an SQL query with a join, a template construct - since this value comes from the same table – or with a parent triples map (for users creating mappings manually) or the linking block (for participants using Juma Uplift). For subtask 3, parent triples map or the linking block would be the expected solution.

The task performance, as it is defined in this paper, is the number of correct triples found in the RDF output generated from the participants’ mappings. Note that the performance takes the output of the mapping into account and not the mapping itself, as there are multiple possible correct solutions, but only one correct output. The Jena API[11] was used to compare the RDF models and count the triples.

Table 1 shows the challenges associated to the task.

Table 1. Challenges associated to the task

Subtask	Short description	Challenge/Non-trivial aspects
#1	Map and type entities to a class with three attributes	One attribute mapping is the concatenation of other two attributes. This requires mapping using a SQL query, the use of a template construct or the data transformation function ‘concatenating’ - for participants using Juma Uplift).
#2	Map and type another entity with one attribute	Map cities as a second entity from the same table using another triples map.
#3	Linking the subjects created in the previous subtasks	Linking subjects created in subtasks 1 and 2. This requires the use of a template construct, a SQL query with a SQL join, the R2RML parent triples map construct for mappings created manually, or the linking block for participants using Juma Uplift.

5 Results and Analysis

In this section, we present the results and analysis of the experiment described in Section 4. As stated in the previous section, in order to test the research hypothesis **H1**, the WP and NASA-TLX instruments were applied. **Table 2** shows the perceived mental workload of both instruments for the R2RML group. **Table 3** shows the same scores for the Juma group.

[11] `https://jena.apache.org/`, accessed May 2018.

Table 2. Perceived mental workload scores for the R2RML group

Participant	WP	NASA-TLX
#1	45.86	65.6
#2	37.86	64.8
#3	41.28	37.8
#4	73.13	51.4
#5	27.86	35.6
#6	32.43	51.6
#7	75.29	56.8
#8	46.29	42
#9	71.13	62.8
#10	16.13	34
#11	58.43	54.4
#12	63.56	56
#13	63.13	73.2
#14	49.56	61.6
AVG	**50.14**	**53.40**
STD	**18.08**	**12.14**

Table 3. Perceived mental workload scores for the Juma group

Participant	WP	NASA-TLX
#1	46.86	52.6
#2	41.29	47
#3	36.57	31.2
#4	54.57	51.2
#5	54.72	48
#6	45.56	61.8
#7	57.87	57.4
#8	46	34.4
#9	54.86	52.4
#10	64.13	48.2
#11	28.43	26.4
#12	43.29	37
AVG	**47.85**	**45.63**
STD	**9.92**	**10.95**

The Anderson-Darling normality test was applied to the R2RML and Juma groups. **Table 4** shows the A values and p-values resulting from this test. **Fig. 4** shows histograms for the same data.

Table 4. Anderson-Darling normality test per group

MWL	R2RML		Juma	
	A	**p-value**	**A**	**p-value**
WP	0.20	0.84	0.23	0.47
NASA-TLX	0.33	0.74	0.39	0.33

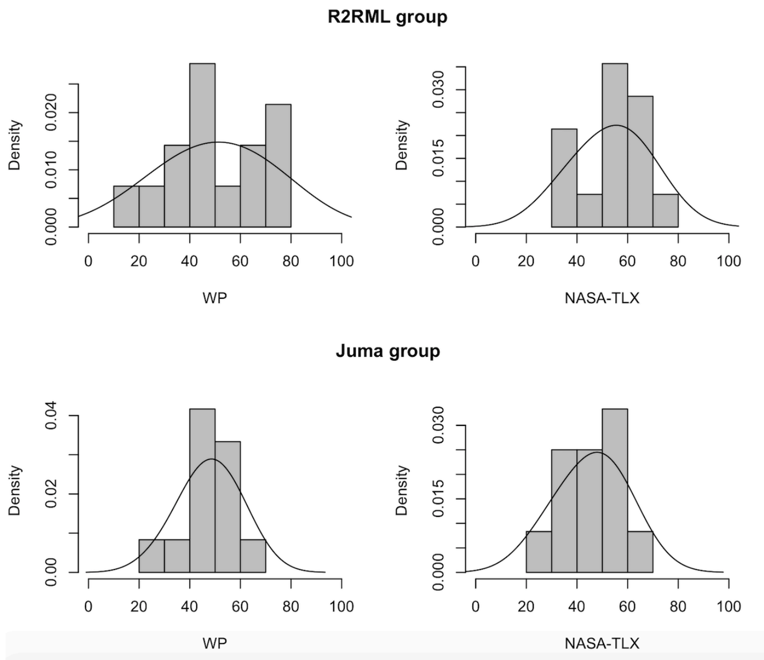

Fig. 4. Mental workload score histograms per group

In order to compare the scores between the groups, we have applied the Welch T-Test and the Wilcoxon test. These tests are used to compare whether two samples are statistically different. The main difference between these tests is that the Welch T-Test assumes normality of the data. The Wilcoxon Test, however, is considered an alternative test when the data does not follow a normal distribution. Considering that the Anderson-Darling test indicates that the data in both groups is normal, the Welch T-Test should be sufficient. For clarity, we have also applied the Wilcoxon test. The results of the independent two sample Welch T-Test and Wilcoxon test are presented in **Table 5**.

Table 5. Mental workload test between groups

MWL	Welch Test		Wilcoxon Test	
	T	**p-value**	**W**	**p-value**
WP	-0.41	0.69	75.5	0.68
NASA-TLX	-1.72	0.10	50	0.08

As mentioned in Section 4.2, the performance of participants was calculated by counting the correct triples in the output of the execution of the mappings created by each participant. In this sense, the R2RML group achieved task performance of 35.98%; while the Juma group achieved 93.08%. **Fig. 5** shows a scatterplot between performance and the MWL scores. In this plot, the correlation between performance and mental workload scores in the R2RML group seems to be multi modal, while the distribution in the Juma group seems to be unimodal. These plots and the smaller standard deviation indicate that the mental workload scored perceived by participants in the Juma group are more consistent than the ones found in the R2RML group.

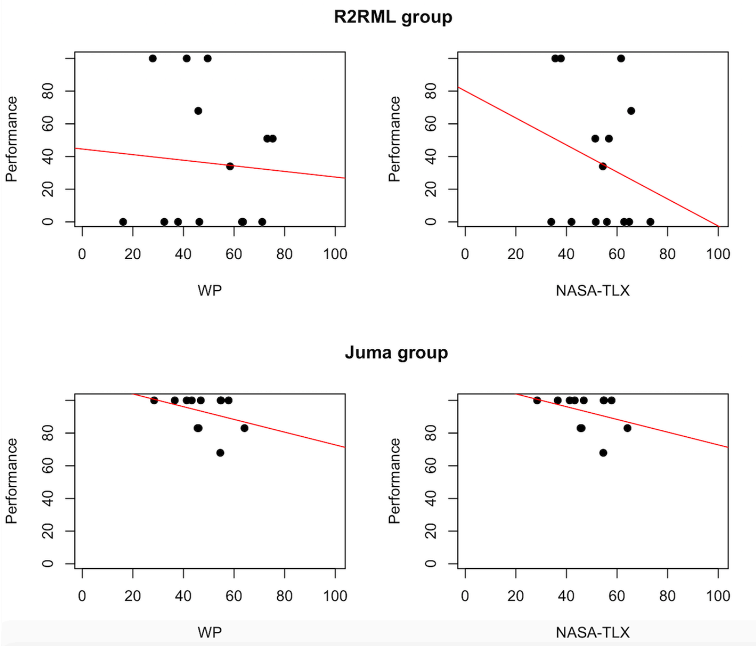

Fig. 5. Scatterplot between MWL scores and performance per group

5.1 Reliability

In order to test the research hypothesis **H2**, the Cronbach's alpha coefficient was applied. Cronbach's alpha is a commonly used measure of reliability within questionnaires. Cronbach’s alpha should coefficients should be higher than 0.70, as it is suggested in the literature [30]. **Table 6** shows the Cronbach’s alphas for the WP and NASA-TLX mental workload instruments. These results highlight a strong internal consistency of the items (questions) in these instruments. They also suggest that these instruments are reliable measures of mental workload.

Table 6. Cronbach's alpha index for WP and NASA-TLX

MWL	Alpha index
WP	0.78
NASA-TLX	0.85

Fig. 6 shows a scatterplot between WP and NASA-TLX scores per group. This plot suggests a positive linear relation between the MWL instruments WP and NASA-TLX. It also indicates that when WP increases, so does the NASA-TLX score.

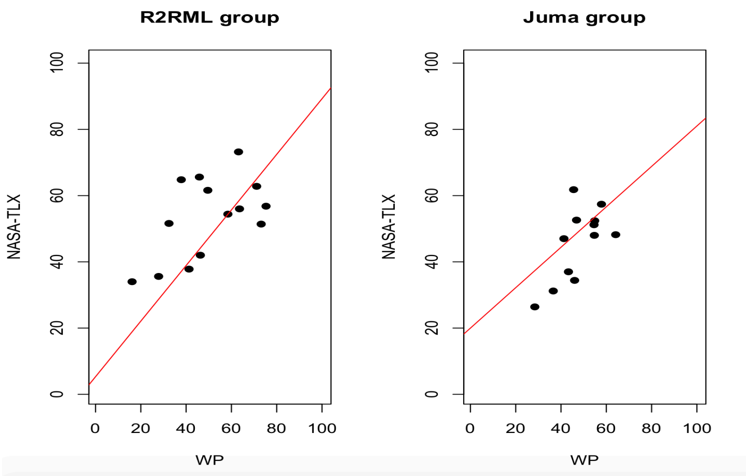

Fig. 6. Scatterplot between the WP and NASA-TLX scores

5.2 Findings

The performance of participants using the Juma representation was higher than for participants manually creating the mappings using R2RML (as per **Fig. 5**). The perceived mental workload scores were slightly smaller for Juma, for the WP and NASA-TLX instruments (**Fig. 4**). It is important to note that the performance achieved by the Juma group is almost three times the performance achieved by the R2RML group, and that the mental workload scores in the Juma group are slightly smaller. The standard deviation in the Juma group is also smaller than the standard deviation found in the R2RML group. This suggests that these mental workload scores are more consistent in the Juma group, which can also be seen in **Fig. 5**. However, the difference between the mental workload scores' groups was found not to be statistically significant, through the independent two sample Welch T-Test and Wilcoxon test, with NASA-TLX presenting the p-value nearest to the threshold of 0.05. Nonetheless, we argue that these results indicate that the hypothesis **H1** is true. However, since the Welch T-Test and Wilcoxon test did not find the differences between the groups to be statistically significant, maybe due to the small sample size, our conclusion is that more experimentation is needed to confirm the hypothesis **H1**.

Cronbach's alpha showed that the MWL through WP and NASA-TLX are reliable instruments for measuring mental workload, thus the research hypothesis **H2** can be accepted and findings reliably considered. **Fig. 6** also suggests evidence for the validity of MWL instruments, showing a high correlation between WP and NASA-TLX scores for both groups, which is expected.

6 Related Work

6.1 Uplift Mapping Representations

Several mappings languages have been proposed in literature. R2RML [9] is the W3C Recommendation mapping language to map relational databases to RDF. Examples of R2RML implementations are db2triples[12], and morph [32]. Sparqlification Mapping Language [40] is another mapping language based on SQL CREATE VIEWS and SPARQL CONSTRUCT queries with support for relational databases and CSV files. SPARQL-Generate [21] is another SPARQL-based mapping language with support for multiple input data formats. A number of tools provide different visual representations for uplift mappings in order to support user engagement. Karma [20] is an example a web-based visual application for uplift mappings where data is loaded before it can be mapped to RDF. Karma presents the ontologies used during the mapping process in a tree structure and the data being mapped as a table. The mapping is represented using a graph. Map-On [38] is another visual web-based editor where the input data and ontologies being mapped are shown as graphs. Assertions between these graphs are used to generate the uplift mapping. Juma [18], as explained in Section 2.3, is a method that uses the block metaphor in the representation of mappings.

6.2 Mental workload applications

Self-assessment measures of MWL include multidimensional approaches such as the NASA's Task Load Index [16], the Subjective Workload Assessment Technique [33], the Workload Profile (WP) [43] as well as unidimensional measures such as the Copper-Harper scale [7], the Rating Scale Mental Effort [48], the Subjective Workload Dominance Technique [45] and the Bedford scale [34]. These procedures have low implementation requirements, low intrusiveness and high subject acceptability. Mental workload assessment is typically conducted to evaluate the cognitive capabilities related to a certain task. This task may be related to operating vehicles [2, 15, 39, 42], user interfaces [23, 24, 26, 27, 37], teaching[35], emergency response [13], amongst others. The NASA-TLX has been used for evaluating user interfaces in health-care [23, 24, 26, 27] or in e-commerce, along with a dual-task objective methodology for investigating the effects on user satisfaction [37]. The NASA-TLX instrument has also been used in an educational context to evaluate teaching methods [35]. Tracy and Albers adopted three different techniques for measuring MWL in web-site design: NASA-TLX, the Sternberg Memory Test and a tapping test [1, 44]. They proposed a technique to identify sub-areas of a web-site in which end-users manifested a higher mental workload during interaction, allowing designers to modify those critical regions. Similarly, [11] investigated how the design of query interfaces influence stress,

[12] `https://github.com/antidot/db2triples`, accessed in May 2018

workload and performance during information search. Here stress was measured by physiological signals and a subjective assessment technique - Short Stress State Questionnaire. Mental workload was assessed using the NASA-TLX and log data was used as objective indicator of performance to characterize search behavior. In [28], the author investigates the relation between usability, mental workload and human performance. A comparison between machine learning techniques used to predict MWL to the NASA-TLX and the Workload Profile instruments is presented in [29]. In the Linked Data domain, MWL instruments have been used to assess ontology visualizations for semantic mappings [14], and exploratory search over Linked Data [17].

As it can be seen in this section, several studies have assessed the mental workload, including in Web systems, such as the work presented in [25], which is the case of the Juma Uplift tool evaluated in this paper. The evaluation of performance and usability of uplift mapping representations can be found in various studies, including for Juma [18]. However, to the author's knowledge, this paper presents the first attempt at evaluating the mental workload of creation and editing uplift mapping representations.

7 Conclusions and Future Work

This study extends the application of MWL instruments by showing how these can be employed for the task of creating uplift mappings in Linked Data. These instruments can guide developers and researchers in creating tools that find the optimal cognitive load on users.

A primary research has been designed and performed to compare the cognitive load of two different approaches that can be used to create uplift mappings. From the many uplift representations available, the W3C-Recommended mapping language to express mappings from relational databases to RDF, R2RML, and Juma, a visual representation for mappings based on the block metaphor, were selected for this study. The experiment presented in this paper separated participants into two groups, one creating mappings "manually" in R2RML, and another using Juma Uplift to create the same mapping. After the time allocated to execute this task, two mental workload instruments were applied to participants, namely the Workload Profile and NASA Task Load Index. Results have shown that participants using Juma Uplift achieved higher performance with slightly smaller, and more consistent, perceived mental workload scores, when compared to participants creating mapping manually. This may suggest that users interact better with the Juma representation, and that it has a smaller learning curve for the task of creating uplift mappings. Cronbach's alpha showed a strong internal consistency of the items of the questionnaires associated to the two selected mental workload instruments, suggesting that these are reliable.

As it was shown in Section 6, uplift mapping representations are commonly evaluated based on the performance and usability of participants, while the mental workload of performing tasks involving these mapping representations is neglected. The findings of this paper show that the cognitive load is a reliable instrument that can be used to compare, and improve, uplift mapping representations.

Future work might include a comprehensive user study to evaluate performance and usability, together with the cognitive load measurements presented in this study, for the task of creating uplift mappings in Linked Data. Future work might also include the evaluation of the interpretability of uplift mapping representations in Linked Data as an additional task performance measure jointly with other self-reporting MWL instruments.

Acknowledgements. This paper was supported by CNPQ, National Counsel of Technological and Scientific Development – Brazil and by the Science Foundation Ireland (Grant 13/RC/2106) as part of the ADAPT Centre for Digital Content Technology (http://www.adaptcentre.ie/) at Trinity College Dublin.

References

1. Albers, M.: Tapping as a Measure of Cognitive Load and Website Usability. Proceedings of the 29th ACM international conference on Design of communication. pp. 25-32 (2011). DOI: 10.1145/2038476.2038481.
2. Balfe, N., Crowley, K., Smith, B., & Longo, L. Estimation of Train Driver Workload: Extracting Taskload Measures from On-Train-Data-Recorders. In International Symposium on Human Mental Workload: Models and Applications (pp. 106-119). Springer, Cham, (2017). DOI: 10.1007/978-3-319-61061-0_7.
3. Bart, A.C., Tibau, J., Kafura, D., Shaffer, C.A., Tilevich, E.: Design and Evaluation of a Block-based Environment with a Data Science Context. IEEE Transactions on Emerging Topics in Computing, 2017. DOI: 10.1109/TETC.2017.2729585.
4. Bizer, C., Heath, T., Berners-Lee, T.: Linked data-the story so far. International journal on semantic web and information systems, 2009, 5(3), pp.1-22. DOI: 10.4018/jswis.2009081901.
5. Cain, B.: A review of the mental workload literature. Tech. rep., Defence Research & Dev.Canada, Human System Integration (2007).
6. Ceriani, M., Bottoni, P.: SparqlBlocks: Using Blocks to Design Structured Linked Data Queries. Journal of Visual Languages and Sentient Systems, 3, pp.1–21, 2017.
7. Cooper, G.E., Harper, R.P.: The use of pilot ratings in the evaluation of aircraft handling qualities. Technical Report AD689722, 567, Advisory Group for Aerospace Research & Development (1969).
8. Crotti Junior, A., Debruyne, C., Brennan, R. O'Sullivan, D.: An evaluation of uplift mapping languages. International Journal of Web Information Systems, 13(4), pp.405-424, 2017. DOI: 10.1108/IJWIS-04-2017-0036.
9. Das, S., Sundara, S., Cyganiak, R.: R2RML: RDB to RDF Mapping Language. (2012) https://www.w3.org/TR/r2rml/.
10. Debruyne, C., O'Sullivan, D.: R2RML-F: Towards Sharing and Executing Domain Logic in R2RML Mappings. In: Workshop on Linked Data on the Web (LDOW 2016).
11. Edwards, A., Kelly, D., Azzopardi, L.: The Impact of Query Interface Design on Stress, Workload and Performance, pp. 691{702. Springer International Publishing, Cham (2015). DOI: 10.1007/978-3-319-16354-3_76.
12. Euzenat, J., Shvaiko, P. Ontology matching (Vol. 18). Heidelberg: Springer, 2007.
13. Fan, J., & Smith, A. P. The impact of workload and fatigue on performance. In International Symposium on Human Mental Workload: Models and Applications (pp. 90-105). Springer, Cham, 2017. DOI: 10.1007/978-3-319-61061-0_6.
14. Fu, B., Noy, N.F., Storey, M.A.: Indented tree or graph? A usability study of ontology visualization techniques in the context of class mapping evaluation. In: International

Semantic Web Conference, ISWC 2013, (pp. 117-134). DOI: 10.1007/978-3-642-41335-3_8.
15. Guastello, S. J., Marra, D. E., Correro, A. N., Michels, M., & Schimmel, H. Elasticity and Rigidity Constructs and Ratings of Subjective Workload for Individuals and Groups. In International Symposium on Human Mental Workload: Models and Applications (pp. 51-76). Springer, Cham, 2017. DOI: 10.1007/978-3-319-61061-0_4.
16. Hart, S.G.: Nasa-task load index (nasa-tlx); 20 years later. In: Human Factors and Ergonomics Society Annual Meeting. vol. 50. Sage Journals (2006). DOI: 10.1177/154193120605000909.
17. Hoefler. P., Granitzer, M., Veas, E.E., Seifert, C.: Linked Data Query Wizard: A Novel Interface for Accessing SPARQL Endpoints. In: Workshop on Linked Data on the Web (LDOW 2014).
18. Junior, A.C., Debruyne, C., O'Sullivan, D.: Using a block metaphor for representing R2RML mappings. In: Proceedings of the 3rd International Workshop on Visualization and Interaction for Ontologies and Linked Data (VOILA@ISWC 2017).
19. Junior, A.C., Debruyne, C., O'Sullivan, D.: Juma Uplift: Using a Block Metaphor for Representing Uplift Mappings. In: 12th IEEE International Conference on Semantic Computing (ICSC 2018). DOI: 10.1109/ICSC.2018.00037.
20. Knoblock, C., Szekely, P., Ambite, J.L., Goel, A., Gupta, S., Lerman, K., Muslea, M., Taheriyan, M., Mallick, P.: Semi-Automatically Mapping Structured Sources into the Semantic Web. In: Proceedings of the 9th Extended Semantic Web Conference (ESWC 2012). DOI: 10.1007/978-3-642-30284-8_32.
21. Lefrançois, M., Zimmermann, A., Bakerally, N. A SPARQL extension for generating RDF from heterogeneous formats. In: European Semantic Web Conference, ESWC 2017, (pp. 35-50). DOI: 10.1007/978-3-319-58068-5_3.
22. Longo, L.: A Defeasible Reasoning Framework for Human Mental Workload Representation and Assessment. *Behaviour & Information Technology*, *34*(8), 758-786 (2015). DOI: 10.1080/0144929X.2015.1015166.
23. Longo, L., Dondio, P.: On the relationship between perception of usability and subjective mental workload of web interfaces. In: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2015, 2015. DOI: 10.1109/WI-IAT.2015.157.
24. Longo, L.: Designing medical interactive systems via assessment of human mental workload. In: 28th IEEE International Symposium on Computer-Based Medical Systems, CBMS 2015, 2015. DOI: 10.1109/CBMS.2015.67.
25. Longo, L.: Formalising human mental workload as non-monotonic concept for adaptive and personalised web-design. In International Conference on User Modeling, Adaptation, and Personalization, pp. 369-373. Springer, Berlin, Heidelberg, 2012. DOI: 10.1007/978-3-642-31454-4_38.
26. Longo, L.: Human-computer interaction and human mental workload: Assessing cognitive engagement in the world wide web. In: Human-Computer Interaction - INTERACT 2011 - 13th IFIP TC 13 International Conference, 2011. DOI: 10.1007/978-3-642-23768-3_43.
27. Longo, L.: Mental workload in medicine: Foundations, applications, open problems, challenges and future perspectives. In: 29th IEEE International Symposium on Computer-Based Medical Systems, CBMS 2016, 2016. DOI: 10.1109/CBMS.2016.36.
28. Longo, L.: Subjective Usability, Mental Workload Assessments and Their Impact on Objective Human Performance. In: Human-Computer Interaction – INTERACT 2017 - 16th IFIP TC 13 International Conference, 2017. DOI: 10.1007/978-3-319-67684-5_13.

29. Moustafa, K., Luz, S., Longo, L.: Assessment of mental workload: A comparison of machine learning methods and subjective assessment techniques. In: Human Mental Workload: Models and Applications - First International Symposium, HWORKLOAD 2017. DOI: 10.1007/978-3-319-61061-0_3.
30. Nunnally, J. C.: Psychometric theory (2nd ed.). New York: McGraw-Hill, 1978.
31. Pinkel, C., Binnig, C., Haase, P., Martin, C., Sengupta, K., Trame, J.: How to Best Find a Partner? An Evaluation of Editing Approaches to Construct R2RML Mappings. In: European Semantic Web Conference (ESWC 2014). DOI: 10.1007/978-3-319-07443-6_45.
32. Priyatna, F., Corcho, O, Sequeda, J.: Formalisation and experiences of R2RML-based SPARQL to SQL query translation using morph. In 23rd International World Wide Web Conference, WWW'14, Seoul, Republic of Korea, 2014, pp. 479–490. DOI: 10.1145/2566486.2567981.
33. Reid, G.B., Nygren, T.E.: The subjective workload assessment technique: A scaling procedure for measuring mental workload. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, Advances in Psychology, vol. 52, chap. 8, pp. 185{218. North-Holland (1988). DOI: 10.1016/S0166-4115(08)62387-0.
34. Rizzo, L., Dondio, P., Delany, S.J., Longo, L.: Modeling Mental Workload Via Rule-Based Expert System: A Comparison with NASA-TLX and Workload Profile, pp. 215-229. Springer International Publishing, Cham (2016). DOI: 10.1007/978-3-319-44944-9_19.
35. Rizzo, L., Longo, L.: Representing and inferring mental workload via defeasible reasoning: A comparison with the NASA task load index and the workload profile. In: Proceedings of the 1st Workshop on Advances In Argumentation In Artificial Intelligence co-located with XVI International Conference of the Italian Association for Artificial Intelligence (AI*IA 2017), 2017.
36. Rubio, S., Diaz, E., Martin, J., Puente, J.M.: Evaluation of subjective mental workload: A comparison of swat, nasa-tlx, and workload profile methods. Applied Psychology 53(1), 61-86 (2004). DOI: 10.1111/j.1464-0597.2004.00161.x.
37. Schmutz, P., Heinz, S., Metrailler, Y., Opwis, K.: Cognitive load in ecommerce applications: Measurement and effects on user satisfaction. Advances in Human- Computer Interaction 2009. DOI: 10.1155/2009/121494.
38. Sicilia, Á., Nemirovski, G., Nolle, A.: Map-on: A Web-Based Editor for Visual Ontology Mapping. Semantic Web Journal, 8(6), pp.969-980 (2017). DOI: 10.3233/SW-160246.
39. Smith, A. P., & Smith, H. N. Workload, fatigue and performance in the rail industry. In International Symposium on Human Mental Workload: Models and Applications (pp. 251-263). Springer, Cham, 2017. DOI: 10.1007/978-3-319-61061-0_17.
40. Stadler, C., Unbehauen, J., Westphal, P., Sherif, M.A., Lehmann, J. Simplified RDB2RDF Mapping. In: Workshop on Linked Data on the Web (LDOW 2015).
41. Thakker, D., Dimitrova, V., Lau, L., Yang-Turner, F., Despotakis, D.: Assisting user browsing over linked data: requirements elicitation with a user study. In International Conference on Web Engineering, 2013 (pp. 376-383). DOI: 10.1007/978-3-642-39200-9_31.
42. Tong, S., Helman, S., Balfe, N., Fowler, C., Delmonte, E., & Hutchins, R. Workload Differences Between On-road and Off-road Manoeuvres for Motorcyclists. In International Symposium on Human Mental Workload: Models and Applications (pp. 239-250). Springer, Cham, 2017. DOI: 10.1007/978-3-319-61061-0_16.
43. Tsang, P.S., Velazquez, V.L.: Diagnosticity and multidimensional subjective workload ratings. Ergonomics 39(3), 358{381 (1996). DOI: 10.1080/00140139608964470.
44. Tsang, P.S.: Mental workload. In: Karwowski, W. (ed.) International Encyclopedia of Ergonomics and Human Factors (2nd ed.), vol. 1, chap. 166. Taylor & Francis (2006).

45. Vidulich, M.A., Ward Frederic G., S.J.: Using the subjective workload dominance (sword) technique for projective workload assessment. Human Factors Society 33(6), 677{691 (1991). DOI: 10.1177/001872089103300605.
46. Wickens, C.D., Hollands, J.G.: Engineering Psychology and Human Performance. Prentice Hall, 3rd edn. (September 1999).
47. Wickens, C.D.: Multiple resources and mental workload. Human Factors 50(2), 449{454 (2008). DOI: 10.1518/001872008X288394.
48. Zijlstra, F.R.H.: Efficiency in work behaviour. Doctoral thesis, Delft University, The Netherlands (1993).

Appendix A: MWL questionnaires

Table 7. The Workload Profile questionnaire

Label	Question
WP_1	How much attention was required for activities like remembering, problem-solving, decision-making, perceiving (detecting, recognizing, identifying objects)?
WP_2	How much attention was required for selecting the proper response channel (manual - keyboard/mouse, or speech - voice) and its execution?
WP_3	How much attention was required for spatial processing (spatially pay attention around)?
WP_4	How much attention was required for verbal material (eg. reading, processing linguistic material, listening to verbal conversations)?
WP_5	How much attention was required for executing the task based on the information visually received (eyes)?
WP_6	How much attention was required for executing the task based on the information auditorily received?
WP_7	How much attention was required for manually respond to the task (eg. keyboard/mouse)?
WP_8	How much attention was required for producing the speech response (eg. engaging in a conversation, talking, answering questions)?

Table 8. The NASA Task Load Index questionnaire

Label	Question
NT_1	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
NT_2	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
NT_3	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
NT_4	How hard did you have to work (mentally and physically) to accomplish your level of performance?
NT_5	How successful do you think you were in accomplishing the goals, of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
NT_6	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?