Learning Analytics and Survey Data Integration in Workload Research

Studentischer Workload: Definition, Messung und Einflüsse
QUANTEL, 23 June 2016

Evgenia Samoilova; Florian Keusch; Tobias Wolbring (School of Social Sciences/University of Mannheim)
Workload as an essential component of teaching effectiveness (KEMBER, 2004; MARSH, 2001)

- fit between student effort and the course task (COPE & STAEHR, 2005)
- excessive workload is associated with surface learning (BACHMAN & BACHMAN, 2006) and lack of success (COPE & STAEHR, 2005)
European Credit Transfer System (ECTS)

- workload as a common “currency”
- workload defined as time an individual student needs to spend on all learning activities within class as well as outside of class (i.e., internship and individual study time)
Problem Definition: Why Workload?

Workload in online learning

- workload as a “control check”: is an online class as demanding as a traditional class?
- workload as a predictor of dropout in online learning (BOWYER, 2012; ASHBY, 2004)

![Graph showing the number of courses and user accounts on the Coursera platform from February 2012 to March 2013.](image)

Coursera (2013)
Learning Analytics (LA)

- includes the measurement, collection, analysis, and reporting of learners’ data.
- Is used for the “purposes of understanding and optimising learning and the environments in which it occurs” (FERGUSON, 2012, p. 305).
- is usually used for predicting disengagement and dropouts, but not yet used for the measurements of workload.

Example of the reported LA at Coursera (available for instructors)
Problem Definition: Why LA?

Survey Data

+ common way of measuring workload
  BUT
  - additional task of recall and estimation
  - could be a sensitive question: social desirability bias
  - nonresponse and missing data

LA Data

+ low burden for respondents
+ no problem with nonresponse
+ less prone to social desirability bias
  BUT
  - new (not well researched) method of measuring workload
  - cannot capture subjective states such as beliefs, attitudes, and satisfaction
  - incomplete
**Problem Definition: Why LA?**

**Survey Data**

- + common way of measuring workload
  
  BUT
  
  - additional task of recall and estimation
  
  - could be a sensitive question: social desirability bias
  
  - nonresponse and missing data

**Measurement Error**

**LA Data**

- + low burden for respondents
  
  + no problem with nonresponse
  
  + less prone to social desirability bias
  
  BUT
  
  - new (not well researched) method of measuring workload
  
  - cannot capture subjective states such as beliefs, attitudes, and satisfaction
  
  - incomplete
Pilot Study: Goals

• Compare two different methods of workload measurement – survey data collection and LA

• Workload defined as time an individual spends on learning activities

• LA is not “Validation data”: it is NOT error-free

• Instead: Criterion Validity (Basic sanity check)

During the past week, how much time did you spend (in hours) on watching pre-recorded lecture-videos?

Survey data

LA
Pilot Study: Methodology

• 12 weeks of an online course “Fundamentals of Survey and Data Science” (February and May 2016)

• Required course within International Program in Survey and Data Science (IPSDS)

• Course Components: pre-recorded video lectures, weekly online assignments, weekly required and recommended readings, and synchronous meetings via the online video conferencing system BlueJeans
Pilot Study: Methodology

- 16 participants (all working professionals)
- median age 29.5
- 10 women, 6 men
- 2 students were located outside of Europe

<table>
<thead>
<tr>
<th></th>
<th>mean/%</th>
<th>median</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working hours (week)</td>
<td>40.75</td>
<td>41</td>
<td>12.19</td>
</tr>
<tr>
<td>First online course</td>
<td>56 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours/week expected to spend on the course</td>
<td>8.69</td>
<td>8</td>
<td>3.02</td>
</tr>
<tr>
<td>Familiarity with the subject taught in the course:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Not at all familiar</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- A little familiar</td>
<td>25 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Somewhat familiar</td>
<td>44 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Very familiar</td>
<td>31 %</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
• was meaningful only for watching pre-recorded video lectures
• collected via *mediasite* software
• students could watch videos only via streaming, no downloading allowed
• videos allowed for pausing, moving forward and backward by jumping to a specific point, rewatching (parts) of the video, and changing the speed of the video

<table>
<thead>
<tr>
<th>Participant</th>
<th>Views</th>
<th>Total time watching</th>
<th>Time covered</th>
<th>% Watched</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>00:14:00</td>
<td>00:10:00</td>
<td>100%</td>
</tr>
</tbody>
</table>

Example for a 10 minute video B
Data Source #2: Survey Data

- 12 weekly web-based surveys programmed in unipark EFS survey software version EFS 10.9

- Questions: time-use including, three items from the ARCS motivation scale by Keller (2009), satisfaction with the learning materials of a week, and perceived level of stress in the respective week

- Survey invitations were sent to students every Friday evening after the deadline for the submission of the weekly assignment
<table>
<thead>
<tr>
<th>Time (in hours) spent on watching videos (all 12 weeks)</th>
<th>Mean</th>
<th>Median</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA</td>
<td>1.19</td>
<td>1.32</td>
<td>0.48</td>
</tr>
<tr>
<td>Survey</td>
<td>2.67</td>
<td>2.15</td>
<td>1.32</td>
</tr>
</tbody>
</table>

Other Survey Workload Items

<table>
<thead>
<tr>
<th>Time (in hours) spent on completing assignment</th>
<th>Mean</th>
<th>Median</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doing Readings (both required and recommended)</td>
<td>3.51</td>
<td>2.83</td>
<td>1.57</td>
</tr>
<tr>
<td>Discussing Course topics (outside of weekly online meetings)</td>
<td>0.05</td>
<td>0</td>
<td>0.11</td>
</tr>
<tr>
<td>Other course-related work</td>
<td>0.74</td>
<td>0.6</td>
<td>0.57</td>
</tr>
<tr>
<td>Time (in hours) spent on completing assignment</td>
<td>1.19</td>
<td>1.32</td>
<td>0.48</td>
</tr>
</tbody>
</table>
Video watching (in hours) by the data source used
Video watching (in hours) by the data source used

\[ \text{cor(Survey VW, LA VW)} = 0.16 \]

Survey Data

LA Data
Mean Plot of the survey and LA measures of video watching (in hours)

- Factorial repeated-measures ANOVA: Data Source & Weeks
- Eta-squared: significant modest effect of the data source (0.269, p<0.05)
- Very small effect of the week (0.006, p<0.05).
- Very small effect of the interaction (0.008, p>0.05).
<table>
<thead>
<tr>
<th>Data Source</th>
<th>Grades for weekly assignments</th>
<th>Satisfaction with the weekly units</th>
<th>Self-reported time spend on assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey</td>
<td>0.18</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>LA</td>
<td>0.53</td>
<td>0.07</td>
<td>-0.51</td>
</tr>
</tbody>
</table>
Conclusion

• The two methods not only provide different average estimates for the time spent on watching prerecorded lecture videos (with a difference of 1.5 hours), but also seem to have different relationships with other variables.

• Further investigation is needed to identify what exactly causes the difference (e.g. cognitive interviews).

• In both cases, the average workload is below the designed workload of 12 hours per week. Is average workload appropriate for assessing the effectiveness of the teaching? In the further education for professionals heterogeneity of learners’ background is an issue. An alternative approach could be to create different typologies of the learners and their workload.

• The two data sources provide us with more information than we would have yielded based on LA or the survey data alone.
Did we learn anything new?

**Survey Data**
- + Common way of measuring workload
  
  **BUT**
  - Additional task of recall and estimation
  - Could be a sensitive question: social desirability bias
  - Nonresponse and missing data

**LA Data**
- + Low burden for respondents
- + No problem with nonresponse
- + Less prone to social desirability bias
  
  **BUT**
  - New (not well researched) method of measuring workload
  - cannot capture subjective states such as beliefs, attitudes, and satisfaction
Did we learn anything new?

LA Data

#1 Regulations on data privacy in Germany
#2 A notable time investment at the stage of data management
Thank you!