Assessing the validity of a learning analytics expectation instrument: A multinational study

Citation for published version:
Whitelock-Wainwright, A, Gasevic, D, Tsai, Y-S, Drachsler, H, Scheffel, M, Muñoz-Merino, P, Tammets, K & Delgado Kloos, C 2020, 'Assessing the validity of a learning analytics expectation instrument: A multinational study', *Journal of Computer Assisted Learning*, vol. 36, no. 2, pp. 209-240.
https://doi.org/10.1111/jcal.12401

Digital Object Identifier (DOI):
10.1111/jcal.12401

Abstract

To assist Higher Education Institutions in meeting the challenge of limited student engagement in the implementation of Learning Analytics services, the Questionnaire for Student Expectations of Learning Analytics (QSELA) was developed. This instrument contains 12 items, which are explained by a purported two-factor structure of *Ethical and Privacy Expectations* and *Service Expectations*. As it stands, however, the QSELA has only been validated with students from UK University students, which is problematic on account of the interest in Learning Analytics extending beyond this context. Thus, the aim of the current work was to assess whether the translated QSELA can be validated in three contexts (an Estonian, a Spanish, and a Dutch University). The findings show that the model provided acceptable fits in both the Spanish and Dutch samples, but was not supported in the Estonian student sample. In addition, an assessment of local fit is undertaken for each sample, which provides important points that need to be considered in future work. Finally, a general comparison of expectations across contexts is undertaken, which are discussed in relation to the General Data Protection Regulation (GDPR, 2018).

1. Introduction

Engaging with stakeholders to understand what is expected from Learning Analytics (LA) remains an important undertaking and challenge for Higher Education Institutions (Tsai & Gašević, 2017a). Students, as a stakeholder group, have remained largely absent from LA implementation discussions, which can lead us to question the value and appropriateness of those service that are eventually offered (Tsai et al., 2018). To resolve the issue of unequal stakeholder engagement from the student population, the Questionnaire for Student Expectations of Learning Analytics (QSELA; Authors, 2019) was created. By using the QSELA, Higher Education Institutions have a means of gauging what students expect from LA services, with data being used to influence implementation decisions (e.g., what services to emphasise; Tsai et al., 2018). More importantly, it promotes greater engagement with a stakeholder group whose role in LA decision-making has largely been marginal. Motivation of this current work is to increase the accessibility to the QSELA and promote engagement with stakeholders in LA implementation decisions.

1.1. Expectations of Learning Analytics

The initial items of the QSELA were generated on the basis of four themes that were identified from a review of the LA literature; these were: *Ethical and Privacy Expectations, Agency Expectations, Intervention Expectations,* and *Meaningfulness Expectations* (Authors, 2019; 2017). Each of these four themes are well captured by the items of the QSELA (Authors, 2019) and thereby offers higher education institutions a wide-ranging insight into student expectations of LA services. In order to provide a comprehensive understanding of the theoretical basis for the QSELA, each theme, along with their representative factor, will be discussed in turn.

Discussions relating to the ethical procedures involved in LA service implementations have been extensive. In particular, the work undertaken by Sclater (2016) has played an important role in making higher education institutions aware of privacy and ethical issues associated with the collection and analysis of students' educational data. However, this particular work has been dominated by the inputs of institutional managers, practitioners, and researchers; whereas, student input has been relatively low. Even though the development of a code of practice is fundamental to the establishment of LA services that uphold ethical and privacy concerns (Sclater, 2016), the input from students cannot be overlooked (Aguilar, 2017), particularly with reference to ethical and privacy decisions (Slade & Prinsloo, 2014).

When engaged in discussions regarding potential LA services, students have been found to express discomfort once they are made aware that their educational data is amenable to analysis (Roberts, Howell, Seaman, & Gibson, 2016). Additional work by Ifenthaler and Schumacher (2016) shows that students may in fact be open to the collection of educational data, but draw the line at the use of identifiable data. The importance of engaging students in talk centred on ethical and privacy beliefs is further reinforced in our explorations of student attitudes toward LA practices (Authors, 2018A). In this work, we found that students are open to a higher education institution collecting and analysing data, but only for purposes that are considered to be legitimate (Authors, 2018A). Taken together, these abovementioned points show students to hold beliefs towards the ethical and privacy elements of LA services. In particular, while students may consider it acceptable for a university to collect and analyse specific forms of data, but not when data is identifiable or when data is used for illegitimate purposes.

Existing frameworks attempt to encourage institutions to engage data subjects in the implementation of LA services (Drachsler & Greller, 2016), yet input from students in LA services continues to be limited (Tsai & Gašević, 2017a). With accumulating evidence showing students holding strong beliefs toward the privacy and ethical elements of LA services (Arnold & Sclater, 2017; Ifenthaler & Schumacher, 2016; Roberts et al., 2016; Slade & Prinsloo, 2014) (Authors, 2018A), and the potential ideological gap that may arise following unequal engagement of stakeholders (Tsai & Gasevic, 2017; Authors, 2017), the inclusion of the *Ethical and Privacy Expectations* theme items was considered to be important. Of the 12 retained QSELA items, five items relate to the theme of *Ethical and Privacy Expectations* (Appendix 1), which cover beliefs toward providing consent to third party usage of educational data, whether universities seek additional consent for any further usage of the data, and consenting to use any identifiable data (Authors, 2019). These items were found to load onto a distinct factor titled *Ethical and Privacy Expectations* (Authors, 2019) and thereby increases the level of student engagement in issues of transparency and consent (Sclater, 2016; Slade & Prinsloo, 2014).

The remaining seven items of the QSELA load onto a *Service Expectations* factor, which is composed of items related to the *Agency Expectations, Intervention Expectations,* and *Meaningfulness Expectations* themes (Authors, 2019; 2017). This distinction between *Ethical and Privacy Expectations* and *Service Expectations* is important, as it shows that student beliefs toward LA are not restricted to only ethical and privacy issues, but extends into the types of services they want to receive. Researchers have explored student beliefs toward LA services, but this has been restricted to expectations of dashboard features (Roberts, Howell, & Seaman, 2017; Schumacher & Ifenthaler, 2018). Although important in the development of a specific LA service, dashboards are not the only service that can be offered through an institution's implementation of LA (Pardo, Jovanovic, Dawson, Gašević, & Mirriahi, 2017). The QSELA addresses this particular issue by providing institutions, researchers, and practitioners with an insight into students' general beliefs towards the possible services introduced with LA.

The theme of *Agency Expectations* relates to the central tenant of self-regulated learning, which is the ability of students to make their own choices based on the feedback received from LA services (Winne & Hadwin, 2012). This further relates to the need for student-centred learning analytics, as put forward by Kruse and Pongsajapan (2012). LA viewed through the perspective argued by Kruse and Pongsajapan (2012) suggests that students should be able to make sense of their own data, making reflections on their progress, and using this information to decide whether to change their current learning strategy. It is important for students to remain active agents within their own learning, rather than LA services creating a culture of passivity. The QSELA contains two *Service Expectation* items pertaining to the *Agency Expectations* theme. These items seek to explore student beliefs toward making their own decisions on the basis of LA service feedback (item 7, Appendix 1) and whether teaching staff are obligated to act (item 11, Appendix 1). As stated by Prinsloo and Slade (2017), while a higher education institution holds a moral responsibility to act in situations where a student may underperform, this does not remove the responsibility of a student to learn. LA services are typically associated with the implementation of early interventions to offset the possibility of students failing a course (Campbell, DeBlois, & Oblinger, 2007). Nevertheless, it is important for institutions to be mindful of not removing student independence, but balancing this with a level of awareness of whether any student is at-risk of failing or is underperforming. Results from items 7 and 11 can then provide an important insight into whether the student population expect institutions to make decisions on their behalf, or whether learner agency should be upheld.

The *Intervention Expectations* theme items of the QSELA encompass the regularity of feedback (item 4, Appendix 1), the incorporation of LA input in teacher feedback (item 10, Appendix 1), and the use of feedback to promote academic skill development (item 12, Appendix 1). While the development of early alert systems has come to characterise LA services (Campbell et al., 2007), implemented intervention programmes have fallen short of expectations (Dawson, Jovanovic, Gašević, & Pardo, 2017). However, focusing only on early-alert systems is an overly narrow perspective of LA services, particularly in light of developing tools aimed at facilitating self-regulated learning (Winne & Hadwin, 2013), improving the student-teacher relationship (Liu, Bartimote-Aufflick, Pardo, & Bridgeman, 2017), or improved student reporting systems (Bodily & Verbert, 2017). Although these LA services are advantageous for students, it remains necessary for the perspectives of students to be accommodated into these developments (Ferguson, 2012).

The importance of engaging students in discussions around LA service developments such as dashboards have been recommended (Verbert et al., 2014), and progress is being made (Roberts et al., 2017; Schumacher & Ifenthaler, 2018). More specifically, the work of Schumacher and Ifenthaler (2018) and Roberts et al. (2017) show students to want features that allow students to compare their performance to their peers or the provision of real-time feedback, to name a few. In other words, LA services should not be centred on the inclusion of early alert systems; instead, higher education institutions should be seeking to offer a wider variety of support (Ifenthaler & Schumacher, 2016). Moreover, to ensure that students are satisfied with the LA service implemented, it is necessary for researchers to continue to understand what students expect following the disclosure of personal information, and this extends beyond ethical and privacy discussions. Thus, the items of the QSELA related to the purported *Intervention Expectations* theme can be used to add weight to the abovementioned findings by providing an insight into the features students want from the implemented LA service.

The remaining theme of *Meaningfulness Expectations* refers to the LA services being in a format that is applicable and relevant to students (Authors, 2019). Put differently, positive changes in behaviour following the exposure to LA service feedback is predicated on their perceived utility (Wise, Vytasek, Hausknecht, & Yuting Zhao, 2016). The importance of feedback that is pedagogical meaningful has also been raised by teaching staff, who expressed preference for information that can provide an informative understanding of a student's learning activity (Ali, Hatala, Gašević, & Jovanović, 2012). For students, feedback from LA services needs to promote effective learning (Gašević, Dawson, & Siemens, 2015), as feeding back trivial measures is unlikely to make positive changes to their learning. As outlined by Nicol and Macfarlane-Dick (2006), feedback should provide students with the information they require to understand how to proceed in their learning. In other words, feedback should identify gaps and provide insight into how the student can move from their current learning state to a desired state (Nicol & Macfarlane-Dick, 2006). This form of feedback is therefore facilitating a student's ability to meta-cognitively monitor and subsequently regulate their learning (Winne & Hadwin, 2012). Provision of simple performance measures are unlikely to facilitate such changes in learner behaviour and may not reflect what students want. As identified by Schumacher and Ifenthaler (2018), students expect to receive feedback that facilitates their ability to monitor their learning progress, which reinforces the need to engage students in LA service implementation decisions (Gašević et al., 2015). Without understanding or aligning a LA service with the expectations students hold toward the meaningfulness of feedback, it is unlikely that LA services will be used to their full extent due to the effects of negative disconfirmation (Brown, Venkatesh, & Goyal, 2014). As shown by the work of Schumacher and Ifenthaler (2018), it is necessary to understand what LA service features students expect in order for it to be meaningful to support their learning. Those QSELA items capturing the *Meaningfulness Expectations* theme can add weight to the growing body of work showing that students hold beliefs toward the types of LA service features that could support their learning.

1.2. QSELA Background

Prior work in LA that has used questionnaires to understand student perspectives have adopted a perception based approach (Arnold & Sclater, 2017; Ifenthaler & Schumacher, 2016). Although the findings of these provide an important insight into knowing the views students hold towards LA services, the questionnaires do rely upon students having some experience with LA services. Given the low adoption rates of LA services in Europe (Tsai & Gašević, 2017b), it is unlikely that students have formulated perceptions of LA services. There is little consistency in the delivery of LA services and this makes it difficult for students to thoroughly understand what inclusive of such services. With the introduction of the GDPR (2018) in response to the growing concerns around the use of identifiable data, it is becoming ever important that the student perspective towards such issues is understood. Failing to address students' data concerns may then have consequences for the future willingness of students to utilise LA services. A resolution to this aforementioned approach has been presented by Schumacher and Ifenthaler (2018), which is to measure expectations of LA services. In doing so, it allows researchers and higher education institutions to understand what students expect from LA services, which can then inform future implementation decisions. Moreover, a focus on expectations overcomes the shortcomings of measures that require students to have had previous experience with LA services. Based on these points, the QSELA was grounded in the work on expectations.

Under the QSELA framework, a LA service expectation is defined as a ‘belief about the likelihood that future implementation and running of LA services will possess certain features’ (Authors, 2019). As the term expectation is quite general, it was decomposed on the basis of the work of Thompson and Suñol (1995) into ideal and predicted expectations. These specific forms of expectations refer to what an individual desires (ideal expectation) and what are the conditions students expect in reality (predicted expectation). In other words, while desires reflect an unrealistic expectation, a more realistic expectation of the LA service can also be obtained. Thus, researchers and practitioners who utilise the QSELA can differentiate between those LA service features students would ideally want and those that students believe they are most likely to receive.

The development and validation of the QSELA led to 12-items being retained (Appendix 1), which are explained by a purported two-factor structure (Authors, 2019; Figure 1). These two factors correspond to *Ethical and Privacy Expectations* and *Service Expectations*, which can refer to student beliefs toward the ethical procedures involved in LA services (e.g., the university will obtain consent for the collection and analysis of any educational data) and how they would like to benefit from such services (e.g., students receiving regular updates about their learning progress), respectively. These two constructs are largely supported by the literature from the LA field and from prior work with the student population (Authors, 2018A; (Arnold & Sclater, 2017; Ifenthaler & Schumacher, 2016; Roberts et al., 2017, 2016; Schumacher & Ifenthaler, 2018; Slade & Prinsloo, 2014).

Up to now, the QSELA has only been validated within UK (United Kingdom) higher education institutions. Consequently, this means that the QSELA is restricted in its use as the cross-cultural validation of the instrument has yet to be explored. The current study seeks to address the limitation of the QSELA by investigating whether the original factor structure (see Figure 1, Authors, 2019) can be recovered and validated in three European contexts (Spain, Netherlands, and Estonia). In doing so, this will enable a greater number of institutions to use the QSELA in their pursuit of implementing LA services. More importantly for the field of LA, it will increase the engagement from the student population, meeting the challenge Tsai and Gasevic (2017) identified.

1.3. Current Research

As outlined by Authors (2019), the four themes identified in the LA literature (*Ethical and Privacy Expectations, Agency Expectations, Intervention Expectations,* and *Meaningfulness Expectations*) were used to generate 79 items. These were then subject to peer review and reduced down to 37 items. Reasons for the items being removed were down to them being unclear or content overlap. The remaining items were then piloted using students (n = 210) from a higher education institution. Respondents completed the survey and provided comments on the clarity and understanding of each item. The quantitative results obtained were used in a subject to exploratory factor analysis (EFA) and other scale refinement steps (remove highly correlated items), whilst the qualitative comments were used to make adjustments to the wordings of each item. In addition to using student feedback to alter the wordings of each item, further peer review was undertaken. Following these steps, the 37 items were reduced down to 19 items. As the items had been re-worded and communalities remained low, a further distribution to students (n = 674) at the same higher education institution was undertaken, with the results being subject to EFA. The authors were left with a 12-item instrument, with five items loading onto an *Ethical and Privacy Expectations* factor (items 1, 2, 3, 5, and 6; Appendix 1) and seven items loading onto a *Service Expectations* factors (items 4, 7, 8, 9, 10, 11, and 12; Appendix 1).

The model presented in Figure 1 is the purported factor structure identified through the exploratory analysis stages of the instrument development (Authors, 2019). In order to validate this factor structure using confirmatory factor analysis (CFA) and exploratory structural equation modelling (ESEM), which was used as a confirmatory tool (Marsh, Morin, Parker, & Kaur, 2014), a further sample of students (n = 191) from a different higher education institution completed the 12-item instrument (Authors, 2019). For both the ideal and predicted expectation scales, the findings supported the original two-factor structure of the QSELA (Authors, 2019). In addition, study 3 showed that the subscales (*Ethical and Privacy Expectations* and *Service Expectations*) had good measurement quality across both scales (ideal and predicted expectations). Thus, in the context of UK higher education institutions, the QSELA was found to be both internally consistent and valid.

Insert Figure 1 About Here

Irrespective of the QSELA strengths, it has only been validated in a single language. Given the interest of LA services outside of the UK (Ferguson et al., 2015), it is important that stakeholders are readily engaged in implementation decisions across each context. Up to now, only the work of Arnold and Sclater (2017) tangibly explores student expectations of LA services across different locales (students from UK institutions and students from American institutions), which showed the expectations from students to not be consistent across both contexts. It is important to note, however, that the American students had prior experience with LA services, whilst the UK students did not. Although this exposure to LA services does represent an important confound, Arnold and Sclater's work still represents a step towards understanding the consistency of student expectations cross-culturally.

The nominal amount of research exploring student expectations in more than one cultural context represents a gap in the LA literature that needs to be addressed, particularly with regards to determining the suitability of a one-size fits all approach to LA service policy decisions[1]. Put in a different way, it is necessary to understand whether the expectations students hold are culturally consistent and if not, there is a need to develop institution-specific LA service policies. However, to address this limitation using the QSELA, it first needs to be translated and validated in each context. In undertaking these steps to validate the QSELA, it will provide a greater number of higher education institutions with the means of incorporating the expectations of students into their LA services implementation decisions. On this basis, the aim of the current paper is to determine the validity of the QSELA following its translation for use in three different contexts (i.e., Spain, Estonia, and the Netherlands). In addition, an examination of item descriptive statistics was undertaken as a preliminary step towards understanding whether student expectations of LA services are culturally consistent.

2. Methods

2.1. Samples

2.1.1. Estonia

The translated version of the QSELA was distributed through an online survey system at an Estonian university. A total of 161 responses were received (Females = 137). Students were aged between 19 and 60 (*Mean* = 29.63, *Median* = 27, *SD* = 9.38). Majority of the sample were undergraduates (63%, *n* = 101), 35% of the sample were masters students (*n* = 56), and 2% were PhD students (*n* = 4). Of the sample, 11% were taking a science subject (*n* = 18), 4% were taking an engineering subject (*n* = 7), 38% were studying a social science subject (*n* = 61), 39% were taking an arts and humanities subject (*n* = 62), 2% were studying

[1] http://sheilaproject.eu/

a medicine and health science subject (*n* = 4), and 6% categorised their subject as other (*n* = 9).

2.1.2. Spain

The translated version of the QSELA was distributed through an online system to students from a Spanish university and 543 responses were received (Females = 272). The ages of respondents ranged from 16 to 57 (*Mean* = 21.15, *Median* = 20, *SD* = 5.04). Majority of the sample was composed of undergraduate students (87%, *n* = 470), 12% were master students (*n* = 67), and 1% were PhD students (*n* = 6). Of these students, 45% were studying a subject from social and legal sciences (*n* = 244), 41% were taking an engineering subject (*n* = 224), and 14% were studying a subject from humanities, communication, and documentation (*n* = 75). In terms of student type, 93% of the sample were Spanish (*n* = 507), whilst the remaining students were international students (7%, *n* = 36).

2.1.3. The Netherlands

A total of 1,247 students (Females = 705) from a Dutch university completed the Dutch translated version of the 12-item QSELA (Appendix 15). Seven respondents did not provide their age or gave an incorrect age based on the demographic information of the university (e.g., 99 years of age). Of those respondents that did, their ages ranged from 18 to 82[2] (*Mean* = 44.81, *Median* = 46, *SD* = 12.14). Majority of the sample were undergraduate students (64%, *n* = 793), 36% were masters students (*n* = 450), and 4 were PhD students. Respondents were almost equally distributed across the three faculties at the university, 33% (*n* = 413) from culture and jurisprudence, 33% (*n* = 416) from management, science, and technology, and 34% (*n* = 418) from psychology and education. Majority of the sample were Dutch students (90%, *n* = 1125), 9% were European students (*n* = 106), with only 1% of respondents being overseas students (*n* = 16).

[2] The age range was also checked with the student services of the institution who confirmed the upper age limit of the students was correct.

2.2. Instrument

2.2.1. Estonian Version of the QSELA

The 12-item QSELA was translated into Estonian (Appendix 5) for the purposes of the data collection. The process by which the QSELA was translated involved one researcher initially translating the survey into Estonian. A further researcher then translated the Estonian version back to English and this was then check by other colleagues to understand the meaning conveyed in the items. This enabled the researchers to determine whether the original meaning of the QSELA items were preserved in the translated version. Following these steps, slight amendments were made to the Estonian version of the QSELA in order to align the concepts and terms within the educational system. As with previous distributions (Authors, 2019), responses to the items were made on two 7-point Likert scales (1 = Strongly Disagree; 7 = Strongly Agree) corresponding to ideal (Ideally, I would like that to happen) and predicted (In reality, I would expect that to happen) expectations.

2.2.2. Spanish Version of the QSELA

The original 12-item QSELA was translated into Spanish (Appendix 10) by a researcher who was a native Spanish speaker and who was fluent in English. Once translated, a further researcher assessed the quality of the translation to determine whether the original meaning of the QSELA items had been preserved. If there were any identified discrepancies, the researchers made subtle changes to the translation in order to better align the item wordings with the original QSELA. As with the original instrument, responses were made on two 7-point Likert scales (1 = Strongly Disagree; 7 = Strongly Agree) corresponding to ideal (Ideally, I would like that to happen) and predicted (In reality, I would expect that to happen) expectations.

2.2.3. Dutch Version of the QSELA

The original 12-item QSELA was translated into Dutch (Appendix 15). This was undertaken by a colleague whose is a native Dutch speaker. Once translated, two researchers, who are native Dutch speakers, assessed the translated survey to determine whether the original meaning of the QSELA items had been preserved. If there were any identified discrepancies, the researchers made subtle changes to the translation in order to better align the item wordings with the original QSELA. The translated instrument was then distributed to students through an online survey system. As with the original instrument, responses were made on two 7-point Likert scales (1 = Strongly Disagree; 7 = Strongly Agree) corresponding to ideal (Ideally, I would like that to happen) and predicted (In reality, I would expect that to happen) expectations.

2.3. Analysis Overview

For each sample (Estonian, Spanish, and Dutch students), the raw data was analysed using both CFA and ESEM in Mplus 8 (Muthén & Muthén, 2017), and geomin rotation was used for the ESEM (Asparouhov & Muthén, 2009). Therefore, to avoid reiterating the same analysis details for each sample, this section presents all the details regarding the methodological steps undertaken. This will involve an assessment of response distributions, details regarding how the model fit will be assessed, and how localised sources of strain will be identified.

The decision to analyse the data using both CFA and ESEM was based on the work of Marsh et al. (2014), which questioned the suitability of CFA. This is due to the requirement of zero cross-loadings, which results in instruments that appear ill-fitting (Marsh et al., 2014), and factor correlations that are inflated (Asparouhov & Muthén, 2009). In contrast, ESEM allows for items to cross-load, and can be used as either an exploratory or confirmatory tool (Marsh et al., 2014). Thus, by allowing cross-loadings, ESEM leads to more accurate factor correlation estimates, but also the identification of problematic items (i.e., items with high loadings on the non-target factors). This may then allow for the identification of problems that would go unnoticed when only using CFA.

An examination of the response distributions (Appendices 2 to 4) showed there to be a ceiling effect, particularly in relation to the ideal expectation scale. This was anticipated, as this level of expectation refers to what students' desire from a LA service; thus, representing an upper reference point of the service students want. Due to the presence of this ceiling effect, the scale-shifted approach to the unweighted least squares estimation (ULSMV) was used for both the CFA and ESEM (Muthén, Muthén, & Asparouhov, 2015). This estimator choice was also based upon it being advantageous in small sample sizes, but also yields more accurate parameter estimates when it converges (Forero, Maydeu-Olivares, & Gallardo-Pujol, 2009; Muthén et al., 2015).

To assess the fit of each model, the X^2 test is reported along with the following alternative fit indices: Comparative Fit Index (CFI), Tuker-Lewis Index (TLI), and Root-Mean Square Error of Approximation (RMSEA). These three indices are typically presented as additional means of evaluating a model due to the various issues associated with the X^2 statistic and negligible misspecifications (Bandalos, 2018); as detailed below, these indices have their own limitations. In relation to the alternative fit indices, Hu and Bentler's (1999) suggested cut-offs of .95 for CFI and TLI, and .06 for RMSEA have been regularly used as indicators of good fitting models. Whilst others have suggested that RMSEA values between .08 and .10 to be indicative of a mediocre fit (MacCallum, Browne, & Sugawara, 1996). The problem, however, is that these cut-offs were based on the maximum likelihood estimator, not categorical estimators such as ULSMV. As shown in the work of Xia (2016), it is inappropriate to generalise the Hu and Bentler criteria to occasions when the ULSMV estimator is used due to its dependency upon thresholds. In addition, the simulation study of McNeish, An, and Hancock (2018) has shown these alternative fit indices (i.e., CFI and RMSEA) to be affected by the measurement quality of the model. Specifically, increased standardised factor loadings result in model fit indices that would be indicative of poor fit (Hancock & Mueller, 2011). For McNeish and colleagues, they recommend that evidence of measurement quality should be given in order to provide a context for fit indices (McNeish et al., 2018). Thus, for the CFA the standardised factor loadings will be presented along with the average loading for each scale. In terms of the ESEM, the range and mean absolute factor loadings will be provided.

In the case of a significant X^2 test, an assessment of localised strain within the model is necessary (Kline, 2015; Ropovik, 2015). To do this, an examination of residual correlations is presented (Kline, 2015), in conjunction with modification index (MI) and standardised expected parameter change (SEPC) values (Saris, Satorra, & Veld, 2009). For residual correlations, absolute values ≥ .10 are indicative of localised strains (Kline, 2015). Whereas, MI values ≥ 3.84 (Brown, 2015), in addition to SEPC values ≥ .10 (Saris et al., 2009), point to local misfit within the model. In the event that misfit is identified, it is then important to consider whether a respecification of the model, which allows for correlated errors between the problematic variable pair, is theoretically justified. As shown in the previous work, both scales of the QSELA (ideal and predicted expectations) showed local misfits between items 2 and 5 and items 11 and 12 (Authors, 2019). However, based on the content of these items there was no justification for the respecification of the model that allowed the errors of these aforementioned items to correlate. This evidence was taken into account if the same sources of misfit were found in the current work.

Finally, it is important to note that the ESEM is being used in a confirmatory approach, as recommended by Marsh et al. (2014). Based on prior work, we have proposed a two-factor structure (*Ethical and Privacy Expectations* and *Service Expectations*; Figure 1) that explains students' expectations towards LA services. Thus, there is a defined factor structure that is guiding the current work, which is to validate the QSELA in three contexts (Estonian, Spanish, and Dutch students). In addition, the approach put forward by Marsh et al. (2014), which is to compare the fits from both the CFA and ESEM is followed. According to Marsh and colleagues if, on comparison, the models show differences in fits that are marginal then the results of the more parsimonious CFA model are presented. Analysing the data using a combination of CFA and ESEM was further substantiated on the basis of the immaturity of the research and the sources of previously identified misfit (Authors, 2019).

Descriptive statistics are also presented to provide a general illustration of how the QSELA results can be used to inform LA service implementations. Given the main aim of validating the QSELA in three additional locales, the responses to each item were not discussed in detail. Instead, the approach used was to explore what items received the highest and lowest score, on average, across each sample and scale.

3. Results

3.1. Summary of Results

The originally purported factor structure (*Ethical and Privacy Expectations* and *Service Expectations*) was supported in two of the three contexts (the Netherlands and Spain) for both expectation scales (ideal and predicted). In the case of the Estonian student sample, items 4, 5, and 7 did not load onto their target factor (*Service Expectations*). On this basis, the Estonian version of the QSELA can be questioned on the grounds of validity. It is therefore important for additional research to be undertaken with a larger sample than was used here (n = 161) to determine whether the QSELA is suitable for use in Estonian higher education contexts. As for the Dutch and Spanish versions of the QSELA, the findings support the use of this instrument to measure student expectations of learning analytics services in higher education contexts of these countries.

In addition to an assessment of the factor structure, an exploration of the mean values for the QSELA items obtained from the three countries was undertaken. For the Estonian student sample, respondents had high expectations towards being provided with a learning profile and ensuring all data was secure, but generally expressed indifference towards the obligation to act. High expectations towards data security were also found for both the Dutch and Spanish. As for the *Service Expectations* items, the Spanish students had higher ideal expectations towards receiving regular feedback, but higher predicted expectations for complete learning profiles. In the case of the Dutch student sample, knowing how progress compared to set goals was the *Service Expectation* item with the highest mean across both expectation scales (ideal and predicted).

A detailed presentation of these results by context is presented in the following sections (sections 3.2, 3.3, and 3.4). This covers an evaluation of the validity of each scale in conjunction with an overview of the obtained descriptive statistics. A discussion follows the presentation of the findings obtained from each context with a view of grounding them into the previous literature on student perspectives of learning analytics services. This detailed presentation of the results is also presented in Table 1, which shows how the content is broken into subsections across each sample.

Insert Table 1 About Here

3.2. Detailed Results for Estonian Student Sample

3.2.1. ESEM and CFA Results

Ideal Expectation Scale

The two-factor model, when fitted using ESEM, resulted in an acceptable fit ($X^2(43, n = 161) = 107.42$, $p < .001$, RMSEA = .10 (90% CI .07, .12), CFI = .95, TLI = .93) and was marginally better than the CFA model ($X^2(53, 161) = 145.58$, $p < .001$, RMSEA = .10 (90% CI .08, .13), CFI = .93, TLI = .92; output presented in Appendix 6). Given the marginal improvement obtained by the ESEM, the results of this model are presented.

The ESEM results showed the two factors (*Ethical and Privacy Expectations* and *Service Expectations*) to be strongly correlated (.60). The factor loadings are presented in Table 2, which shows all items to load highly (> .40) on their target factors (i.e., items 1, 2, 3, 5, and 6 load on the *Ethical and Privacy Expectations* factor and items 4, 7, 8, 9, 10, 11, and 12 load on the *Service Expectations* factor). The absolute factor loadings, $|\lambda|$, for the *Ethical* *and Privacy Expectations* factor ranged from .01 to .86, with a mean of .42. Whereas, the $|\lambda|_{\text{Service Expectations}}$ range from 0 to 1.01 (M = .45). Even though the item loadings were stronger for their target factor, there are two cross-loadings that needed to be highlighted. These were for item 11 and item 12, which have cross-loadings of -.30 and -.39 on the *Ethical and Privacy Expectation* factor. However, these loadings remained lower than their target factor loadings (.72 and 1.01 for items 11 and 12, respectively). While the target factor loading of item 12 exceeded 1, this could be found when factors are correlated (Jöreskog, 1999).

Although the alternative fit indices were suggestive of an acceptable fit, the X^2 test was found to be significant; thus, an inspection of local fit was warranted (Kline, 2015; Ropovik, 2015). Starting with the modification indices and standardised expected parameter change values, there were two possible modifications to be made by freely estimating the correlated errors between items 7 and 8 (MI = 10.19, SEPC = .37) and items 11 and 12 (MI = 18.47, SEPC = .61). An assessment of the absolute correlation residual values (Appendix 7) provided further evidence of localised strain between these items, with values of .12 (between items 7 and 8) and .13 (between items 11 and 12). Previous work on this scale (Authors, 2019) identified localised strain within the purported two-factor structure, specifically between items 11 and 12. As discussed within this prior work, there is no justification for modifying the model to permit correlated errors between items 11 and 12. With regards to the misfit between items 7 and 8, this has not been previously identified, but from a content perspective there is no justification for a re-specification that allows the errors of these items to correlate.

Insert Table 2 About Here

Predicted Expectation Scale

An improved model fit was obtained using ESEM (X^2(43, *n* = 161) = 118.05, *p* < .001, RMSEA = .10 (90% CI .08, .13), CFI = .97, TLI = .95) compared to the CFA (X^2(53, *n* =161) = 197.79, *p* < .001, RMSEA = .13 (90% CI .11, .15), CFI = .94, TLI = .93; output presented in Appendix 8). As the ESEM resulted in a better fitting model, the results of this are reported.

The results of the ESEM showed the two factors to strongly correlate (.62), with all items strongly loading (> .40) onto their target factors (items 1, 2, 3, 5, and 6 on the *Ethical and Privacy Expectations* factor, and items 4, 7, 8, 9, 10, 11, and 12 on the *Service Expectations* factor; Table 3). More specifically, $|\lambda|_{\text{Ethical and Privacy Expectations}}$ ranged from .01 to .92 (*M* = .41) and $|\lambda|_{\text{Service Expectations}}$ ranged from 0 to .93 (*M* = .47). While majority of the items loaded highly onto their target factors, there were some cross-loadings that were suggestive of possible misspecifications. For instance, item 4 had a loading of .43 on factor two (*Service Expectations*) and a loading of .40 on factor one (*Ethical and Privacy Expectations*). Based on the content of the item (receiving regular updates based on the analysis of any educational data) and prior work (Authors, 2019), item 4 was not expected to cross-load. Although not to the same degree as item 4, both item 5 and item 7 showed cross-loadings that could also be problematic (.34 and .33, respectively). Taken together, the ESEM had identified a number of misspecifications related to item loadings, which required further investigations.

Adding to the abovementioned misspecifications related item loadings, an examination of local fit further pointed to additional model problems. An assessment of the residual correlations showed there to be five absolute values that were ≥ .10 (Appendix 9), these items were also found to have large MI and SEPC values. The specific sources of misfit were between items 3 and 4 (-.10; MI = 12.90, SEPC = -.56), items 7 and 8 (.12; MI = 17.21, SEPC = .45), items 8 and 9 (.10; MI = 13.44, SEPC = .42), items 9 and 11 (-.11; MI = 17.26, SEPC = -.41), and items 11 and 12 (.12; MI = 26.06, SEPC = .64). The only misspecification that had previously been identified was between items 11 and 12 and it was stated that the correlation between these errors could not justified. Similarly, correlating the errors of other items that have been identified (e.g., items 7 and 8) could be supported on conceptual grounds.

Insert Table 3 About Here

3.2.2. Descriptive Statistics

Table 4 presents the means and standard deviations for each item across expectation types (ideal and predicted) for the Estonian student sample. Based on a comparison of mean values for each expectation type, the average responses were always higher for the ideal expectation scale than the predicted expectation scale. This adds weight to the ability of the QSELA to differentiate between expectation types (ideal and predicted).

For those items related to the originally proposed *Ethical and Privacy Expectations* factor (items 1, 2, 3, 5, and 6). The highest average response on both the ideal (M = 6.41, SD = 1.12) and predicted (M = 5.86, SD = 1.29) expectation scales was for item 2 (*the university will ensure that all my educational data will be kept securely*). Whereas, the lowest average response on both the ideal (M = 5.81, SD = 1.41) and predicted (M = 5.05, SD = 1.57) expectation scales was for item 5 (*the university will ask for my consent to collect, use, and analyse any of my educational data (e.g., grades, attendance, and virtual learning environment accesses*).

In terms of the *Service Expectations* factor (items 4, 7, 8, 9, 10, 11, and 12), item 9 was both the highest average ideal (M = 5.93, SD = 1.23) and predicted (M = 5.16, SD = 1.36) expectation item. Item 9 stated that *the learning analytics service will present me with a complete profile of my learning across every module (e.g., number of accesses to online material and attendance*. Item 11 however, had the lowest average response for both ideal (M = 5.29, SD = 1.73) and predicted (M = 4.09, SD = 1.73) expectation types. The content of item 11 was: *the teaching staff will have an obligation to act (i.e., support me) if the analytics show that I am at-risk of failing, underperforming, or if I could improve my learning*).

Insert Table 4 About Here

3.2.3. Discussion

While the alternative fit indices for both scales (ideal and predicted) show the two-factor model to have acceptable fit, the X^2 test remains significant, and there were a number of misspecifications that could not be ignored. For the ideal expectation scale, while items 11 and 12 loaded highly onto the target factor (*Service Expectations*), they showed weak cross-loadings onto the *Ethical and Privacy Expectations* factor. On the predicted expectation scale, however, item 4 showed a weak factor loading on both the target factor (*Service Expectations*) and non-target factor (*Ethical and Privacy Expectations*). In addition to item 4, items 5 and 7 also showed weak cross-loadings onto their non-target factors (*Service Expectations* and *Ethical and Privacy Expectations*, respectively). Thus, based on these points it is clear that the Estonian version of the QSELA, based on the current sample, did not provide support for the purported two-factor model. Given the small sample size (*n* =161), it remains necessary that further work is undertaken to assess the validity of the Estonian QSELA using larger samples. In addition, the current work has adopted a confirmatory approach in the use of ESEM, which has identified weaknesses in applying the two-factor structure to the Estonian context. The next step may be for researchers to undertaken an exploratory approach to assess whether a refinement in the items is needed or whether an alternative factor structure can be proposed.

Based on the descriptive statistics presented in Table 4, a general view of what the sample of Estonian students expect from LA services is given. From an ethical and privacy perspective, they have strong expectations regarding the maintenance of security over any data collected. Whereas, the belief that consent should be sought before educational data is collected and analysed did elicit agreement from students, the expectation was not as strong as when compared to ensuring that all data is held securely. It may be that students were open to the university collecting data for legitimate purposes (Authors, 2018A), but concerns over who has access to the collected data resulted in stronger expectations toward data security (Ifenthaler & Schumacher, 2016; Roberts et al., 2016).

The expectations toward the LA service features showed that for the Estonian student sample, they hold stronger beliefs toward receiving a learning profile. Whilst their expectations regarding the implementation of early alert systems was one of indifference. The work of Schumacher and Ifenthaler (2018) has found students to expect LA service features that updated them about their learning progress. These views have also been expressed in the work of Roberts et al. (2016), but here the students were also concerned about the loss of independence on account of the LA service being in place. Taking these aforementioned points into consideration, the findings are suggestive of students considering feedback from LA services as an important supplement to their learning, as it could allow students to evaluate their progress toward a set goal (Winne & Hadwin, 2012). Whereas, the possibility of early-alert systems may undermine the agency that students exercise whilst they learn (Kruse & Pongsajapan, 2012), and LA should not remove responsibility a student has to learn (Prinsloo & Slade, 2017). This further reinforces the importance of understanding what students expect from LA services (Ferguson, 2012), as it is clear that while higher education institutions may consider some features to be useful (e.g., early-alert systems), it may not coincide with student expectations.

3.3. Detailed Results for Spanish Student Sample

3.3.1. ESEM and CFA Results

Ideal Expectations Scale

A marginally improved fit was obtained from the CFA (X^2(53, *n* = 543) = 115.92, *p* < .001, RMSEA = .05 (90% CI .04, .06), CFI = .98, TLI = .97) compared to the ESEM (X^2(43, *n* = 543) = 109.74, *p* < .001, RMSEA = .05 (90% CI .05, .07), CFI = .97, TLI = .96; output present in Appendix 11). As the CFA model was more parsimonious, the results from this model are presented.

The unstandardised and standardised estimates for the two-factor solution are presented in Table 5. The unstandardised estimates were all statistically significant (*ps* < .001), with a mean standardised loading of .76. The R^2 values showed the two factors to explain a moderate to large amount of the latent continuous response variance (R^2 range = .38 - .66). Both factors (*Ethical and Privacy Expectations* and *Service Expectations*) were found to strongly correlate (.53), but the correlation was at a value that did not suggest poor discriminant validity (i.e., values exceeding .85; Brown, 2015). Moreover, the average variance extracted for the *Ethical and Privacy Expectations* (.55) and *Service Expectations* (.59) factors exceeds the squared of the correlation between the two factors (.28; Fornell & Larcker, 1981).

On the basis of alternative fit indices, the two-factor model could be regarded as having an acceptable fit, but an assessment of local fit was required due to the significant X^2 test (Kline, 2015; Ropovik, 2015). There were only two absolute residual correlation values ≥ .10 (Appendix 12), which were between items 2 and 5 (-.10) and items 11 and 12 (.14). MI and SEPC values also showed that the model fit could be improved by allowing the errors between items 2 and 5 (MI = 12.34, SEPC = -.36) and items 11 and 12 (MI = 27.35, SEPC = .41) to be correlated. These two sources of local misfit within the model had previously been identified (Authors, 2019), but there was no justification for allowing the errors of these items to correlate. Therefore, no modifications to the model were undertaken.

Insert Table 5 About Here

Predicted Expectation Scale

A comparison between the results obtained from both the ESEM (X^2(43, *n* = 543) = 327.78, *p* < .001, RMSEA = .11 (90% CI .10, .12), CFI = .96, TLI = .94; output presented in Appendix 13) and CFA (X^2(53, *n* = 543) = 376.13, *p* < .001, RMSEA = .11 (90% CI.10, .12), CFI = .95, TLI = .94) showed the fits to be marginally different. Thus, a decision was made to report the results of the parsimonious CFA model.

Table 6 shows all unstandardised and standardised estimates from the two-factor structure. All unstandardised estimates were statistically significant (*ps* < .001), with a mean standardised loading of .80. The R^2 values showed the two factors to account for a large amount of the latent continuous response variance (R^2 range = .54-.75). Whilst the two factors were strongly correlated (.70), this correlation did not exceed what would be considered as poor discriminant validity (i.e., .85; Brown, 2015). In addition, the average variance extracted for both factors (.62 and .66 for *Ethical and Privacy Expectations* and *Service Expectations*, respectively) exceeded the square of the correlation (.49; Fornell & Larcker, 1981).

An assessment of the residual correlations (Appendix 14) showed four absolute values that are ≥ .10, which were between items 2 and 3 (.10), items 2 and 12 (.10), items 4 and 5 (.16), and items 8 and 9 (.11). MI and SEPC values were also indicative of misspecifications between items 2 and 3 (MI = 30.45, SEPC = .31), items 2 and 12 (MI = 26.04, SEPC = .31), items 4 and 5 (MI = 66.31, SEPC = .53), and items 8 and 9 (MI = 33.06, SEPC = .44). Whilst the misfit between items 2 and 3 had previously been identified by Authors (2019), the remaining sources of localised strain had not. In either case, there was no justification to re-fit the model with correlated errors between the aforementioned variable pairs.

Insert Table 6 About Here

3.3.2. Descriptive Statistics

Table 7 shows the descriptive statistics for the Spanish student sample across both expectation types (ideal and predicted). Based on a comparison of mean values, it can be seen that average responses on the ideal expectation scale were higher than the predicted expectation scale. Thus, as found with the Estonian student sample, the validity of the QSELA to differentiate between ideal and predicted expectation types is further supported.

Considering only the *Ethical and Privacy Expectation* items (items 1, 2, 3, 5, and 6), the descriptive statistics were similar to those of the Estonian student sample on both expectation types (ideal and predicted). The highest ideal (*M* = 6.61, *SD* = 1.02) and predicted (*M* = 5.64, *SD* = 1.36) expectation mean values were for item 2 (*the university will ensure that all my educational data will be kept securely*). Whereas, the lowest ideal (*M* = 6.01, *SD* = 1.40) and predicted (*M* = 4.67, *SD* = 1.72) expectation mean values were for item 5 – *the university will ask for my consent to collect, use, and analyse any of my educational data (e.g., grades, attendance, and virtual learning environment accesses*.

Whilst the highest and lowest average responses for the *Ethical and Privacy Expectation* items (items 1, 2, 3, 5, and 6) were the same across Estonian and Spanish student samples, there were slight differences with regards to *Service Expectation* items (items 4, 7, 8, 9, 10, 11, and 12). For the Spanish student sample, item 4 (*the university will regularly update me about my learning progress based on the analysis of my educational data*) received the highest average ideal expectation (*M* = 6.17, *SD* = 1.27). Whereas, item 9 received the highest average predicted expectation response (*M* = 5.00, *SD* = 1.73). Item 9 asked to students regarding the following statement: *the learning analytics service will present me with a complete profile of my learning across every module (e.g., number of accesses to online material and attendance*. Although the highest predicted expectation, item 9 received the lowest average response on the ideal expectation scale (*M* = 5.91, *SD* = 1.44). Similar to the Estonian student sample, item 11 (had the lowest average response for the predicted expectation scale (*M* = 4.16, *SD* = 1.81). Item 11 asked whether *the teaching* *staff will have an obligation to act (i.e., support me) if the analytics show that I am at-risk of failing, underperforming, or if I could improve my learning*.

Insert Table 7 About Here

3.3.3. Discussion

The alternative fit indices for the ideal expectation scale would suggest a good fitting model; whereas, the predicted scale fit could only be considered as acceptable. In order to provide a context for these fit indices, an assessment of measurement quality was also provided. This showed the mean standardised loading to be higher on the predicted expectation scale (M = .80) than the ideal expectation scale (M = .76). Thus, from a position of measurement quality, the predicted expectation scale exceeded that of the ideal expectation scale.

For both scales, the X^2 test was found to be significant; thus, an inspection of local fit was warranted. In terms of the ideal expectation scale, the sources of misfit (between items 2 and 5 and items 11 and 12) had previously been identified (Authors, 2019). As stated in this prior work, while these items were to some extent related, there was no justified reason for respecifying the model to allow the errors of these items to correlate. Therefore, no steps were taken in the current study to freely correlate the item errors. A different set of localised strains for the predicted expectation scale were identified, with only a single variable pair being previously identified (misfit between items 2 and 3). In none of these cases was there a justifiable reason for respecifying the model with correlated errors between the problematic variable pairs. Taken together, it could therefore be shown that both scales showed good measurement quality, with the predicted expectation scale exceeding that of the ideal expectation scale, and the fit for each scale can at least be considered as acceptable. Nevertheless, further work on the scale is needed, particularly as the X^2 test was found to be significant.

An inspection of those descriptive statistics relating to the *Ethical Privacy Expectations* factor (Table 7) show the expectations of the Spanish student sample to be similar to those held by Estonian student sample. Put differently, as with the Estonian student sample, the Spanish student sample held stronger expectations, on average, toward the university ensuring all data was secure than the university seeking consent before collecting and analysing educational data. This again reiterates the view that students may be more open to their data being used for legitimate purposes (Authors, 2018A), as universities regularly use such data for assessments and to monitor academic progress. Irrespective of these beliefs regarding the provision of consent for the collection and use of educational data, these Spanish students expected the university to ensure that any collected data remains secure (Ifenthaler & Schumacher, 2016; Roberts et al., 2016).

For the items of the *Service Expectations* factor, the Spanish student sample appeared to hold strong ideal expectations towards receiving regular feedback, but had higher predicted expectation towards the provision of complete learning profiles. Similar to Estonian student sample, the Spanish students were seemingly indifferent to the provision of early-alert systems. Again this overview of the descriptive statistics does suggest that features aimed at supporting learner agency and self-regulated learning are expected from LA services (Schumacher & Ifenthaler, 2018). Whereas, early interventions may have unintended consequences (e.g., added pressure for students) or may even be a hindrance to independent learning (Roberts et al., 2016). These concerns could be attributed to the indifference that students expressed towards the possibility of incorporating early-alert systems in LA services.

3.4. Detailed Results for the Dutch Student Sample

3.4.1. ESEM and CFA Results

Ideal Expectation Scale

An improved fit was obtained from the ESEM (X^2(43, *n* = 1247) = 166.63, *p* < .001, RMSEA = .05 (90% CI .04, .06), CFI = .98, TLI = .97) than the CFA (X^2(53, *n* = 1247) = 288.05, *p* < .001, RMSEA = .06 (90% CI .05, .07), CFI = .96, TLI = .95; output presented in Appendix 16). Thus, the results of the ESEM are presented.

The results of the ESEM showed the two factors to weakly correlate (.09), with all items loaded strongly (> .40) onto their target factors (items 1, 2, 3, 5, and 6 on the *Ethical and Privacy Expectations* factor, and items 4, 7, 8, 9, 10, 11, and 12 on the *Service Expectations* factor; Table 8). The $|\lambda|_{\text{Ethical and Privacy Expectations}}$ ranged from 0 to .81 (*M* = .36) and the $|\lambda|_{\text{Service Expectations}}$ ranged from 0 to .90 (*M* = .51). There were no problematic cross-loadings, but item 11 did show a weak cross-loading onto the *Ethical and Privacy Expectation* factor (λ = -.20).

An assessment of local strain in the model was required due to the significant X^2 test (Kline, 2015; Ropovik, 2015). From an inspection of the residual correlation values (Appendix 17), there was only one absolute value ≥ .10, which was between items 11 and 12 (.12). MI and SEPC values also pointed to a possible misspecification between items 11 and 12 (MI = 66.13, SEPC = .42). As previously stated, this misfit within the model had been identified beforehand (Authors, 2019); however, there was no justified reason for allowing the errors of these items to correlate.

Insert Table 8 About Here

Predicted Expectation Scale

A marginal improvement in model fit was obtained using the ESEM (X^2(43, *n* = 1247) = 513.51, p < .001, RMSEA = .09 (90% CI .09, .10), CFI = .96, TLI = .93; output presented in Appendix 18) compared to the CFA (X^2(53, *n* = 1247) = 612.15, p < .001, RMSEA = .09 (90% CI = .09, .10), CFI = .95, TLI = .94). Therefore, the CFA model results are presented on the basis of it being a more parsimonious model.

Table 9 presents both the standardised and unstandardised estimates for the two-factor solution. All unstandardised estimates were statistically significant (*ps* < .001), with a mean standardised loading of .81. The R^2 values showed the two factors to account for a large amount of the latent continuous response variance (R^2 range .42-.79). The two factors were moderately correlated (.43), which did not suggest poor discriminant validity (i.e., did not exceed .85; Brown, 2015). In addition, the average variance extracted for the *Ethical and Privacy Expectations* factor (.69) and the *Service Expectations* factor (.63) exceeded the square of the correlation (.18).

An inspection of the residual correlations (Appendix 19) showed that there were eight instances of absolute values that were ≥.10. Majority of these large residual correlations were for item 11, specifically between item 1 (-.12), item 2 (-.13), item 3 (-.10), and item 12 (.13). MI and SEPC values provided further evidence of misspecification between items 1 and 11 (MI = 42.49, SEPC = -.26), items 2 and 11 (MI = 46.29, SEPC = -.30), items 3 and 11 (MI = 30.76, SEPC = -.29), and items 11 and 12 (MI = 59.39, SEPC = .38). Again, the misfit between items 11 and 12 had been identified, but there are no grounds for respecification (Authors, 2019). The remaining sources of local strain (between item 11 and items 1, 2, and 3) had not been found before; thus, no respecifications of the model was made, but these instances of misfit are further explored. The remaining sources of strain within the model, based on absolute residual correlation values, were between items 1 and 2 (.12; MI = 55.20, SEPC = .44), items 1 and 9 (-.10; MI = 31.13, SEPC = -.28), items 2 and 9 (-.11; MI = 32.25, SEPC = -.32), and items 4 and 5 (.18; MI = 97.86, SEPC = .54). Of these localised areas of strain, only the poor prediction between items 4 and 5 ha been identified previously (predicted expectation scale for the Spanish student sample) and there was no justification for correlated errors. For the remaining variable pairs, there are no grounds for respecifying the model.

Insert Table 9 About Here

3.4.2. Descriptive Statistics

Table 10 presents the mean and standard deviations for each item of the QSELA for the Dutch student sample across expectation types (ideal and predicted). For all items, apart from item 11, the average response was always higher for ideal than predicted expectations. Item 11 asked whether *the teaching staff will have an obligation to act (i.e., support me) if the analytics show that I am at-risk of failing, underperforming, or if I could improve my learning*. An examination of item 11 for the Dutch sample showed that whilst the average responses were similar (*M* = 4.25 and *M* = 4.27 for ideal and predicted expectations, respectively), the standard deviation value for the ideal expectation was the largest across all items (*SD* = 2.06). Thus, for the Dutch student sample there was much variability in regards to their ideal beliefs toward teaching staff having an obligation to act under circumstances where a student may be at-risk of failing. An examination of the response distributions in Appendix 4 shows that for item 11 there were high concentrations of responses for response categories 1 (strongly disagree) and 7 (strongly agree) on the ideal expectation scale. In other words, this item appeared to split student opinions, with a large proportion either expressing a desire of such features or not. Whereas, in the case of the predicted expectation scale, there was a larger proportion of students who expressed indifference to this expectation. It was therefore clear that some students did not desire such services aimed at early interventions, but may express indifference as to whether it will become an actuality. Other than this discrepancy, the descriptive statistics were largely supportive of the Dutch translated version of the QSELA differentiating between ideal and predicted expectations.

Considering only the *Ethical and Privacy Expectation* items, the highest ideal (*M* = 6.69, *SD* = .74) and predicted (*M* = 5.93, *SD* = 1.39) expectations, on average, was for item 2 (*the university will ensure that all my educational data will be kept securely*). Whereas, the lowest average ideal (*M* = 6.21, *SD* = 1.21) and predicted (*M* = 5.38, *SD* = 1.58) expectations was for item 5 – *the university will ask for my consent to collect, use, and analyse any of my educational data (e.g., grades, attendance, and virtual learning environment accesses*.

For the *Service Expectation* items, item 8 (*the learning analytics service will show how my learning progress compares to my learning goals/the course objectives*) received the highest average response on both the ideal (*M* = 5.50, *SD* = 1.67) and predicted (*M* = 5.14, *SD* = 1.54) expectation scales. Similar to the findings from the Estonian student sample, item 11 received the lowest average response on both the ideal (*M* = 4.25, *SD* = 2.06) and predicted (*M* = 4.27, *SD* = 1.66) expectation scales. Item 11 asked whether *the teaching staff will have an obligation to act (i.e., support me) if the analytics show that I am at-risk of failing, underperforming, or if I could improve my learning.*

Insert Table 10 About Here

3.4.3. Discussion

The alternative fit indices obtained from the ideal expectation scale showed the two-factor structure to have a good fit. Moreover, the improved fit was obtained from using the ESEM than the CFA. While the factor loadings presented in Table 8 show all items to load highly (> .40) onto their target factors, item 11 had a small but non-zero negative loading (λ = -.20) on the *Ethical and Privacy Expectations* factor; which was the largest cross-loading.

For the predicted expectation scale, the CFA model was retained due to the differences with the ESEM being marginal. While the alternative fit indices for the two-factor model were found to be acceptable, and the measurement quality was good (mean standardised loading = .81). An assessment of local fit showed there to be a number of strains in the model, particularly related to item 11. Based on the content of these variable pairs (i.e., item 11 with items 1, 2, 3, and 12), there was no justifiable reason for the respecification of the model to include correlated errors. However, focusing only on local strains between item 11 and those variables attributed to the *Ethical and Privacy Expectations* factor (items 1, 2, and 3), there may be other reasons for this misfit. While not presented, the ESEM results for the predicted expectation scale showed item 11 to have a weak negative cross-loading onto the *Ethical and Privacy Expectation* factor (λ = -.18; Appendix 18). Taken together, it is clear that while item 11 is strongly related to the type of service students will receive, specifically whether early interventions should be implemented, there is also an ethical element. As discussed by Prinsloo and Slade (2017), a higher education institution does share some responsibility in relation to the obligation to act, particularly from a moral basis. Thus, this may explain why item 11 weakly cross-loaded onto the *Ethical and Privacy Expectations* factor for both ideal and predicted expectation scales. In other words, students may expect that an ethical LA service would entail a right to decide if teaching staff have an obligation to act if they are deemed to be underperforming or at-risk of failing.

An assessment of local fit in the model did identify a source of strain between the variable pair of items 11 and 12, which had been identified previously (Authors, 2019). Whilst this variable pair been the most frequent source of misfit within the model, it has remained inconsistent. As shown in the Spanish student sample, the misfit between this variable pair (items 11 and 12) was only found for the ideal expectation scale; whereas, this localised strain occurred for both scales (ideal and predicted) in the Dutch and Estonian student samples. Thus, respecification of the two-factor model that included a correlated error between items 11 and 12 could not be justified on theoretical grounds, but also due the inconsistency of this misfit.

Taking the abovementioned points into consideration, it is clear that the ideal expectation scale, based on alternative fit indices, exhibited good fit and all items loaded strongly onto their target factors, with cross-loadings being relatively small. The predicted expectation scale showed an acceptable fit, based on alternative fit indices, but the measurement quality was good. Irrespective of these findings, the X^2 test remained significant for both scales. Whilst an examination of local misfit did not highlight any variable pairs within the model whose errors could be justifiably be correlated, it remains pertinent that researchers continue to assess the validity of the Dutch version of the QSELA.

Based on the descriptive statistics provided in Table 10, similarities with the Spanish and Estonian student samples can be found. In terms of the *Ethical and Privacy Expectations* factor items, the Dutch student sample appear to have strong ideal and predicted expectations toward the university ensuring that all collected data remains secure. Whereas, the weakest item, on average, for both the ideal and predicted expectation scales was for the university obtaining consent for the collection and analysis of educational data. This again shows that students may in fact be open to the university collecting and analysing specific educational data if the purpose is deemed legitimate (Authors, 2018A). However, students hold stronger beliefs toward the university ensuring all collected data remain secure (Ifenthaler & Schumacher, 2016; Roberts et al., 2016).

For the *Service Expectations* factor, the highest mean value on both scales (ideal and predicted) was for students receiving feedback on how their learning is progressing in relation to a set goal. In contrast, the lowest average expectation for both scales (ideal and predicted) was for the provision of an early-alert system. As with the Estonian and Spanish student sample, these descriptive statistics are suggestive of students expecting features that aim to support the regulation of their learning (Schumacher & Ifenthaler, 2018), but remain indifferent to those features that could undermine learner agency (Roberts et al., 2016).

4. Comparing Expectations

The following section seeks to compare the three samples of students based on the means of the 12 QSELA items. In doing so, it provides a general insight into whether the expectations of students are consistent across countries and higher education institutions. This has important connotations for the development of policies that regulate learning analytics services, specifically in determining whether a one size fits all policy or institution-specific policy is more suited[3].

4.1. Comparisons

Figure 2 presents the mean value of each item of the QSELA by country and expectation type (ideal and predicted). What can be taken away from this figure is that students across all samples seemingly have higher expectations (ideal and predicted) toward the *Ethical and Privacy Expectations* factor items (items 1, 2, 3, 5, and 6). In particular, the expectation toward the university ensuring that all data is kept secure (item 2) has the largest mean value across all items on both scales. Whereas, the expectation that the university will seek consent to collect and analyse educational data (item 5) is lowest across each country. In the case of those items related to *Service Expectations* (items 4, 7, 8, 9, 10, 11, and 12), the Spanish student sample to generally have higher expectations, on average, compared to the Estonian and Dutch student samples on the ideal expectation scale. Whereas, the mean values for Dutch student sample on the ideal expectation scale show them to have lower expectations of LA service features. In relation to the predicted expectation scale, the average responses to the items of the *Service Expectations* factor are generally lower than

[3] http://sheilaproject.eu/

responses on the ideal expectation scale. It can also be seen that item 11 receives the lowest average response for each sample.

Insert Figure 2 About Here

4.2. Discussion

Using the descriptive statistics alone, preliminary insights into possible differences in student expectations of LA services can be made, as shown in Authors (2019). With regards to *Ethical and Privacy Expectations*, item 2 (*the university will ensure that all my educational data will be kept securely*) received the highest average response on both the ideal and predicted expectation scales across each sample (Estonian, Spanish, and Dutch students). This is similar to what was found with the sample of UK university students (Authors, 2019). Likewise, item 5, stating that *the university will ask for my consent to collect, use, and analyse any of my educational data (e.g., grades, attendance, and virtual learning environment accesses*, received the lowest average responses on both scales (ideal and predicted expectations) across each sample (Estonian, Spanish, and Dutch students), which was again found by Authors (2019).

From comparing highest and lowest average responses for both ideal and predicted expectation scales on the *Ethical and Service Expectation* items, there is indication of similarities across the different samples. Students hold strong beliefs toward the university securely holding all collected data (item 2), whilst the belief that a university should seek consent before the collection, use, and analysis of educational data appears to elicit the lowest average response for each sample of students (item 5). Although for the ideal expectation scale, the average responses are indicative of students strongly agreeing to item 5. For predictive expectations, responses to item 5 generally show students to be between indifference and weakly agreeing. A plausible assumption here is that it is common place for universities to collect large amounts of educational data in order to evaluate attendance and to contact students; therefore, it may be that students expect such practices to be undertaken without their consent. On the other hand, ensuring that all data remains secure may elicit higher expectations on account of students' personal data being stored by the higher education institution. Thus, whilst educational data is collected by a university, students believe that procedures should be in place that uphold privacy and confidentiality (Ifenthaler & Schumacher, 2016; Roberts et al., 2016; Slade & Prinsloo, 2014).

In relation to the *Service Expectation* items, the descriptive statistics do show variability in what features students expect from LA services. Our prior work with UK university students (Authors, 2019) showed that their highest average ideal expectation response was for item 10 (*the teaching staff will be competent in incorporating analytics into the feedback and support they provide to me*), whilst for predicted expectations this was the lowest average response. The highest average predicted expectation response was for item 8 (*the learning analytics service will show how my learning progress compares to my learning goals/the course objectives*), whilst the lowest average ideal expectation response was for item 11 stating *the teaching staff will have an obligation to act (i.e., support me) if the analytics show that I am at-risk of failing, underperforming, or if I could improve my learning*.

For the Estonian student sample, they held high ideal and predicted expectations of wanting a LA service that provided them with a complete profile of their learning (item 9). As with the UK student sample (Authors, 2019), the Estonian student sample had low ideal expectations toward teaching staff having an obligation to act (item 11), and this was also the lowest predicted expectation item. Likewise, the Dutch student sample was found to have the lowest average response on item 11 for both ideal and predicted expectation scales. Their (the Dutch student sample) highest average responses for both ideal and predicted expectations, however, were for LA services that show students how their learning progress compares to a set goal (item 8). In terms of the Spanish student sample, receiving a complete profile of their learning (item 9) had the lowest ideal expectation on average, but also the highest average response on the predicted expectation scale. Whereas, the highest average response for the ideal expectation scale was for receiving regular updates about their (the students') learning (item 4), and the lowest average response for the predicted expectation was for the belief about teaching staff having an obligation to act (item 11).

It appears that students do not hold strong expectations toward the use of early interventions if LA services found them to be at-risk. Similarly, Roberts et al. (2016) found students to express concern over LA services removing the ability of students to make their own independent decisions. Given the importance placed on independent learning at universities, having systems in place that are centred on the implementation of early interventions to assist underperforming or at-risk students is a contradiction to this position. In line with the view of being independent learners, students appeared to hold higher expectations of LA services that offer informative profiles about their learning, how learning is progressing with reference to a set goal, or receiving regular updates about their learning progress. Thus, students seemingly prefer an LA service that facilitates independent learning rather than one which would impede their self-determination (Schumacher & Ifenthaler, 2018).

5. General Discussion

Equally engaging with stakeholders in the implementation of LA services continues to be a challenge for Higher Education Institutions (Tsai et al., 2018). Failure to accommodate the voice of relevant stakeholders into such discussions may result in a LA service that is not valued and, therefore, falls short of supporting student learning. To help Higher Education Institutions understand what is generally expected from stakeholders, specifically students, the QSELA was developed (Authors, 2019). The aim of this current work was to increase the access to this descriptive questionnaire by undertaking an evaluation of three translated versions (Dutch, Estonian, and Spanish).

It has been shown that the QSELA provides a basic understanding of what students expect from LA services and the possible cross-cultural differences that need to be explored further. In particular, it provides an important stakeholder perspective of what students want from LA services, which is one focused on upholding independence and ensuring that all data is protected. This adds weight to the findings of Roberts et al. (2016), which found students to view LA services as potentially undermining their ability to self-direct their own learning. As discussed by Kruse and Pongsajapan (2012), LA services that predominately focus on interventions may result in a culture of passivity. Rather, students should be provided with feedback that can motivate positive changes to their learning (Gašević et al., 2015), such as engaging in self-regulation (Winne & Hadwin, 2012). What is more, features aimed at promoting more effective learning is what students expect from LA services (Roberts et al., 2017; Schumacher & Ifenthaler, 2018). Thus, the aforementioned points further reinforce the importance of gauging the expectations of students towards the LA service they want, rather than providing a service we believe they want.

The average responses to the *Ethical and Privacy Expectations* items provides an important perspective from the end-users of LA services, particularly in understanding their beliefs towards data handling procedures. Given the new General Data Protection Regulation (GDPR, 2018) that will be put into force in Europe in May 2018, European universities will be required to apply new regulations when it comes to handling student data. These will provide fundamental rights towards the data subject and the data they leave behind. Examples of these rights include: general requirements about transparency and communication, meaningful information about the algorithms involved, information about profiling, access to and rectification of personal data, and the right to erasure (Drachsler & Greller, 2016; Hoel, Griffiths, & Chen, 2017). In other words, universities will be expected to meet the *Ethical and Privacy Expectations* of the QSELA. From a student perspective, we can see that, on average, they have strong ideal expectations toward the university ensuring all data remains secure or controlling thee access from third party companies. However, responses to the predicted expectation scale show students' beliefs to not be as strong. Therefore, while it is desirable for the university to follow such data handling procedures (e.g., asking for consent to use identifiable data), students may not expect too much from their universities, even though the GDPR 2018 demands these. The reason for these lowered predicted expectations may be the result of students may not feel that the university is capable of undertaking such steps when it comes to handling data.

It is also concerning that most students have low expectations of their teaching staff being able to incorporate analytics into the feedback they receive (item 10) or to intervene in circumstances of underperformance (item 11), particularly as the GDPR 2018 requires European Universities to provide a clear purpose for their use of LA services. The latter is referred to as legitimate interests and provides a lawful basis for processing data that prevents an overreliance on obtaining consent (Sclater, 2018). These legitimate interests may encompass the interests of the higher education institution, third party interests, or even societal interests. However, for higher education institutions to rely upon legitimate interests they must balance their interests against those of the students. If students would not expect their data to be used in a particular way, then the interests of the higher education institution become overruled. Applying this in the current context, it can be seen that students do not have high expectations regarding certain elements of a LA service (analytics in teacher feedback and the obligation to act); therefore, a higher education institution could not depend upon legitimate interests to process data for these purposes. This is in line with the points raised by Sclater (2018) that state that higher education institutions cannot use legitimate interests as a basis for carrying out interventions and must instead obtain consent. Taken together, the student perspective towards LA interventions seemingly supports the view that such practices cannot be undertaken without first obtaining consent as it is not an interest for all students.

5.1. Findings Across Cultures

Even though the QSELA is an advantageous instrument to guide LA service implementations, it had so far only been tested in UK higher education institutions (Authors, 2019). The current work sought to address this limitation by validating the three translated versions (Estonian, Spanish, and Dutch) of the QSELA. In doing so, this will increase the number of countries who are able to use the QSELA in their pursuit of implementing LA services. Of the three samples (Etonian, Spanish, and Dutch students) used in this study, the findings from the Estonian student sample are not supportive of the purported two-factor model. Whereas, the results obtained from the Spanish and Dutch student samples show the translated versions of the QSELA to have acceptable fit (based on alternative fit indices) and good measurement quality.

The problems with the Estonian version of the QSELA can be attributed to the cross-loadings that were identified through the use of ESEM (Asparouhov & Muthén, 2009). Whilst four items showed weak cross-loadings onto their non-target factors (i.e., items 5, 7, 11, and 12), item 4 loaded weakly onto both the target and secondary factor (λ = .43 and .40, respectively) for the predicted expectation scale. Given that the current work utilised a confirmatory approach, no respecifications of the model were undertaken in order to address these problematic loadings. Nevertheless, our results highlight strains within the model that require further investigation. The next steps should then be to reassess the Estonian version of the QSELA utilising a larger sample of students. In addition, an exploratory approach to ESEM should be undertaken as items may need to be removed or an alternative factor structure may be proposed. If continued problems are identified, it would show the QSELA to be an inappropriate tool to be used and an alternative instrument may be required.

As for the findings obtained from the Spanish and Dutch student samples, the two-factor structure was supported. If the cut-offs proposed by Hu and Bentler (1999) are used to assess the fit, then the ideal expectation scale appears to provide a better fit. Whilst, the RMSEA values obtained for the predicted expectation scale would be considered as acceptable or poor (MacCallum et al., 1996). As recommended by McNeish et al. (2018), alternative fit indices need to be interpreted within the context of measurement quality, particularly as it is attributed to RMSEA functioning differently. Thus, from a measurement quality, the predicted expectation scale was good, even exceeding the ideal expectation scale.

Irrespective of these results pertaining to alternative fit indices and measurement quality, the X^2 test was significant for each scale and sample (Spanish and Dutch students). It was therefore imperative to conduct an inspection of local fit, paying particular attention to the absolute residual correlation values and both MI and SEPC values. From this assessment of local fit, a number of problematic variable pairs were identified. In none of these cases did a source of misfit lead to a model respecification, which was a decision informed by both prior work (Authors, 2019) and item content. For example, the misfit between items 11 and 12 had been previously identified and it was identified in all three student samples, but not all scales. More specifically, it was not identified for the predicted expectation scale for the Spanish student sample, but was found in the Dutch and Estonian student samples. Therefore, respecifying the model to allow correlated errors between these variables may equate to a capitalisation on chance (MacCallum, Roznowski, & Necowitz, 1992), in addition to there being no justifiable reason (i.e., no overlapping content) for such modifications. Nevertheless, the significant X^2 test shows that further work on the translated versions of the QSELA are required. It may be that an exploratory approach needs to be adopted to understand whether an alternative factor structure needs to be proposed or whether items need to be dropped.

Preliminary insights into possible differences in student expectations have also been reported. For *Ethical and Privacy Expectations*, there appeared to be similarities across the three samples (Estonian, Spanish, and Dutch students). In particular, the descriptive statistics show that, on average, students hold stronger beliefs toward the university ensuring that all data is secure (item 2) over the university seeking consent to collect and analyse educational data (item 5). In the qualitative work with students, Roberts et al. (2016) have found students to express concerns regarding the privacy of their data, particularly in relation to who has access. Similarly, Ifenthaler and Schumacher (2016) found concerns about their privacy to be an important determinant in the acceptance of potential LA services. Taken together, it appears that while students may hold particularly strong beliefs toward providing consent, the institution preserving their privacy is a pivotal expectation.

In regards to *Service Expectations*, students across all three samples seemingly expressed indifference to early interventions (item 11). Whereas, the highest average responses on these items were for LA service features that gave regular updates on their learning (item 4), showed how their learning progress compares to a goal (item 8), or receiving a complete profile of their learning (item 9). As shown by Schumacher and Ifenthaler (2018), students expect LA service features that facilitate self-regulated learning such as being able to monitor their progress. Whilst in the case of early interventions, it has also been found that students express concern when considering how LA services could remove their sense of independence (Roberts et al., 2016). Put in a different way, early interventions do not allow for students to take self-directed actions as it represents an external locus of control (Zimmerman, 2000). Therefore, students may not expect such a service on account of how it may undermine learner agency.

When considering how students respond to the QSELA items across each of the three samples, it is also important to explore possible cultural factors and how they could be factored into future research. In the case of the Dutch university, this was a distance learning institution with a sample that is, on average, older than the either the Estonian or Spanish student samples. Compared to the other samples, the Dutch students appeared to have lower ideal expectations when responding to those items covering service features. Students who are undertaking distance education course need to be self-determined, particularly as they commonly face challenges around time management and family commitments (Nichols, 2010). It may be that the service feature items contained within the QSELA do not align with what such students want, which may include features directed at addressing the loneliness felt in distance students (Croft, Dalton, & Grant, 2010). Beyond the schooling culture of the contexts, there may also be possible economic, social, or political reasons as to why differences in expectations were identified; future research is needed to explore these factors in the context of student expectations of LA services.

5.2. Limitations and Future Research

The findings of the current work raise questions about the suitability of the QSELA in the Estonian context. Given the identified problems regarding cross-loading items, it is important for researchers to follow-up this study with one that adopts an exploratory approach in conjunction with a larger sample size. It may be found that items need to be removed, an alternative factor structure is proposed, or that the QSELA is not a viable instrument to be used in this context. If the latter position is supported, then we encourage researchers to take steps to develop and validate an alternative instrument to measure student expectations of LA services.

For the Spanish and Dutch translated versions of the QSELA, the alternative fit indices do show the model fit to be acceptable. Whilst the RMSEA is high for the Spanish predicted expectation scale, the measurement quality is good and this is associated with the RMSEA functioning (McNeish et al., 2018). Thus, on the basis of these findings it does support the use of the QSELA to measure student expectations within these contexts. Researchers should not be complacent, however, as the X^2 test was significant in all cases and localised strains in the model were identified. Continued assessment of the QSELA in these contexts and beyond should therefore be undertaken.

The QSELA is composed of 12 items that measure student expectations towards the service and ethical elements of LA implementations. It is important to recognise that this is not an exhaustive coverage of all possible features that could form a LA service. Moreover, the items are phrased at a general level as opposed to exploring expectations towards elements of a specific LA service. It is important to note that the QSELA purpose is to provide both researchers and higher education institutions with a means of easily understanding what students expect from LA services in general. This means that it is aiming to understand whether focus should be placed on learner agency and the introduction of early-alert systems, and how data processing should be handled. If researchers do seek to understand student expectations towards a specific LA service then there is a need to develop and validate a new instrument for this purpose.

The provision of descriptive statistics in this study was to provide a preliminary understanding of whether student expectations of LA services are consistent across the three contexts (Estonia, Spain, and the Netherlands). This is, however, limited as there is a need to establish whether the indicators are measuring the same constructs in each group (Kline, 2015). Future research should then seek to assess the measurement invariance of the QSELA cross-culturally and evaluate whether the groups differ in regards to the factors (*Ethical and Privacy Expectations* and *Service Expectations*). In addition to not establishing measurement invariance, the current approach used in this study assumes that students within each of the three groups (Estonia, Spain, and the Netherlands) hold homogenous expectations towards LA services, which is unlikely to be the case. Instead, it can be expected that there is a degree of heterogeneity in student expectations of LA services (i.e., some students may have inflated expectations, whilst other students may express indifference). In order to explore these differences within the student population, future research could segment students into subgroups on the basis of their expectations towards LA services using techniques such as latent class analysis or hierarchical agglomerative clustering. In doing so, it will provide higher education institutions with a nuanced understanding of student expectations of LA services, which could be used to implement LA services that reflect the diversity in what students expects.

A further avenue for future research is to understand how responses obtained from the QSELA compare to and complement those obtained through the use of focus groups, interviews or other qualitative data collection instruments. Qualitative methods could be used as a follow-up to administering the instrument as a means of understanding the reasons why students may have responded in the way they did. Following either of these mixed methods approaches would provide higher education institutions with a detailed understanding of expectation formation and pre-implementation beliefs not captured by the QSELA.

References

Aguilar, S. J. (2017). Learning Analytics: At the Nexus of Big Data, Digital Innovation, and Social Justice in Education. *TechTrends*, 1–9. https://doi.org/10.1007/s11528-017-0226-9

Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education*, *58*(1), 470–489. https://doi.org/10.1016/j.compedu.2011.08.030

Arnold, K. E., & Sclater, N. (2017). Student Perceptions of Their Privacy in Leaning Analytics Applications. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 66–69. https://doi.org/10.1145/3027385.3027392

Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Bandalos, D. L. (2018). *Measurement Theory and Applications for the Social Sciences*. Retrieved from http://ebookcentral.proquest.com/lib/monash/detail.action?docID=5188172

Bodily, R., & Verbert, K. (2017). *Trends and issues in student-facing learning analytics reporting systems research*. 309–318. https://doi.org/10.1145/3027385.3027403

Brown, S. A., Venkatesh, V., & Goyal, S. (2014). Expectation Confirmation in Information Systems Research: A Test of Six Competing Models. *Mis Quarterly*, *38*(3), 729–756.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (Second Edition). New York: The Guilford Press.

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review*, *42*(4), 40–57.

Croft, N., Dalton, A., & Grant, M. (2010). Overcoming Isolation in Distance Learning: Building a Learning Community through Time and Space. *Journal for Education in the Built Environment*, *5*(1), 27–64. https://doi.org/10.11120/jebe.2010.05010027

Dawson, S., Jovanovic, J., Gašević, D., & Pardo, A. (2017). From prediction to impact: Evaluation of a learning analytics retention program. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 474–478. https://doi.org/10.1145/3027385.3027405

Drachsler, H., & Greller, W. (2016). Privacy and Analytics – it's a DELICATE issue. A Checklist to establish trusted Learning Analytics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 89–98. Retrieved from http://dspace.ou.nl/handle/1820/6381

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304. https://doi.org/10.1504/IJTEL.2012.051816

Ferguson, R., Cooper, A., Drachsler, H., Kismihók, G., Boyer, A., Tammets, K., & Monés, A. M. (2015). Learning analytics: European perspectives. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, 69–72. ACM.

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor Analysis with Ordinal Indicators: A Monte Carlo Study Comparing DWLS and ULS Estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 625–641. https://doi.org/10.1080/10705510903203573

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71.

Hancock, G. R., & Mueller, R. O. (2011). The Reliability Paradox in Assessing Structural Relations Within Covariance Structure Models. *Educational and Psychological Measurement*, *71*(2), 306–324. https://doi.org/10.1177/0013164410384856

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, *64*(5), 923–938. https://doi.org/10.1007/s11423-016-9477-y

Jöreskog, K. G. (1999). *How Large Can a Standardized Coefficient be?* Retrieved from http://www.ssicentral.com/lisrel/techdocs/HowLargeCanaStandardizedCoefficientbe.pdf

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford Publications.

Kruse, A., & Pongsajapan, R. (2012). Student-centered learning analytics. *CNDLS Thought Papers*, 1–9.

Liu, D. Y.-T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-Driven Personalization of Student Learning Support in Higher Education. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends* (Vol. 94, pp. 143–169). Retrieved from https://link-springer-com.liverpool.idm.oclc.org/chapter/10.1007/978-3-319-52977-6_5

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490–504. https://doi.org/10.1037/0033-2909.111.3.490

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*, 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

McNeish, D., An, J., & Hancock, G. R. (2018). The thorny relation between measurement quality and fit index cutoffs in latent variable models. *Journal of Personality Assessment*, *100*(1), 43–52.

Muthén, B. O., Muthén, L. K., & Asparouhov, T. (2015). *Estimator choices with categorical outcomes*.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (Eighth Edition). Los Angeles, CA: Muthén & Muthén.

Nichols, M. (2010). Student perceptions of support services and the influence of targeted interventions on retention in distance education. *Distance Education*, *31*(1), 93–113. https://doi.org/10.1080/01587911003725048

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2017). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, 1–11. https://doi.org/10.1111/bjet.12592

Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room: The obligation to act. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 46–55. https://doi.org/10.1145/3027385.3027406

Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give Me a Customizable Dashboard: Personalized Learning Analytics Dashboards in Higher Education. *Technology, Knowledge and Learning*, *22*(3), 317–333. https://doi.org/10.1007/s10758-017-9316-1

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016). Student Attitudes toward Learning Analytics in Higher Education: "The Fitbit Version of the Learning World." *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01959

Ropovik, I. (2015). A cautionary note on testing latent variable models. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01715

Saris, W. E., Satorra, A., & Veld, W. M. van der. (2009). Testing Structural Equation Models or Detection of Misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 561–582. https://doi.org/10.1080/10705510903203433

Schumacher, C., & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior*, *78*, 397–407. https://doi.org/10.1016/j.chb.2017.06.030

Sclater, N. (2016). Developing a code of practice for learning analytics. *Journal of Learning Analytics*, *3*(1), 16–42. https://doi.org/10.18608/jla.2016.31.3

Sclater, N. (2018). Learning analytics and GDPR: What you need to know. Retrieved October 12, 2018, from Jisc website: https://www.jisc.ac.uk/blog/learning-analytics-and-gdpr-what-you-need-to-know-20-sep-2018

Slade, S., & Prinsloo, P. (2014). Student perspectives on the use of their data: Between intrusion, surveillance and care. *European Distance and E-Learning Network*, *18*(1), 291–300.

Thompson, A. G., & Suñol, R. (1995). Expectations as determinants of patient satisfaction: Concepts, theory and evidence. *International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care*, *7*(2), 127–141.

Tsai, Y.-S., & Gašević, D. (2017a). Learning Analytics in Higher Education — Challenges and Policies: A Review of Eight Learning Analytics Policies. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 233–242. https://doi.org/10.1145/3027385.3027400

Tsai, Y.-S., & Gašević, D. (2017b). The State of Learning Analytics in Europe – Executive Summary – SHEILA. Retrieved July 1, 2017, from http://sheilaproject.eu/2017/04/18/the-state-of-learning-analytics-in-europe-executive-summary/

Tsai, Y.-S., Gaševic, D., Whitelock-Wainwright, A., Munoz-Merino, P. J., Moreno-Marcos, Pedro M, Fernandez, A. R., … Kollom, K. (2018). *SHEILA final research report – SHEILA*. Retrieved from https://sheilaproject.eu/2018/11/30/sheila-final-research-report/

Verbert, K., Govaerts, S., Duval, E., Santos, J., Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal & Ubiquitous Computing*, *18*(6), 1499–1514. https://doi.org/10.1007/s00779-013-0751-2

Whitelock-Wainwright, A., Gašević, D., & Tejeiro, R. (2017). What Do Students Want?: Towards an Instrument for Students' Evaluation of Quality of Learning Analytics Services. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 368–372. https://doi.org/10.1145/3027385.3027419

Winne, P., H., & Hadwin, A., F. (2012). The Weave of Motivation and Self-Regulated Learning. In D. Schunk H. & B. Zimmerman J., *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (pp. 297–314). New York: Routledge.

Winne, P. H., & Hadwin, A. F. (2013). nStudy: Tracing and Supporting Self-Regulated Learning in the Internet. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition* *and Learning Technologies* (Vol. 28, pp. 293–308). Retrieved from http://link.springer.com/10.1007/978-1-4419-5546-3

Wise, A. F., Vytasek, J. M., Hausknecht, S., & Yuting Zhao. (2016). Developing Learning Analytics Design Knowledge in the “Middle Space”: The Student Tuning Model and Align Design Framework for Learning Analytics Use. *Online Learning*, *20*(2), 48–75.

Xia, Y. (2016). *Investigating the chi-square-based model-fit indexes for WLSMV and ULSMV estimators* (PhD Thesis). The Florida State University.

Zimmerman, B. J. (2000). Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, *25*(1), 82–91. https://doi.org/10.1006/ceps.1999.1016

Authors (2018A). The privacy paradox in learning analytics – a case study.

Authors (2019). The Questionnaire of Student Expectations of Learning Analytics Questionnaire (QSELA).