Higher Education Studies; Vol. 14, No. 3; 2024
ISSN 1925-4741 E-ISSN 1925-475X
Published by Canadian Center of Science and Education

Insights in Flexible Assessment from Students' and Teachers' Perspectives: A Focus Group Study

Norbert G. C. Huyer[1, 2], Jeroen Dikken[1, 3, 5], Ellen Sjoer[2], R. I. (Vana) Hutter[6, 7], Anne Venema[8] & Peter G. Renden[1, 4]

[1] Department of Nursing, Faculty of Health, Nutrition & Sport, The Hague University of Applied Sciences, The Hague, The Netherlands

[2] Research Group Sustainable Talent Development, Centre of Expertise Global and Inclusive Learning, The Hague University of Applied Sciences, The Hague, The Netherlands

[3] Research Group Urban Ageing, Centre of Expertise Governance of Urban Transitions, The Hague University of Applied Sciences, The Hague, The Netherlands

[4] Research Group Relational Care, Centre of Expertise Health Innovation, The Hague University of Applied Sciences, The Hague, The Netherlands

[f] Research Group Rehabilitation and Technology, Centre of Expertise Health Innovation, The Hague University of Applied Sciences, The Hague, The Netherlands

[6] Netherlands Institute for the Study of Crime and Law Enforcement Amsterdam, the Netherlands

[7] Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, the Netherlands

[8] Educational & Research Policy Department, Inholland University of Applied Sciences, The Hague, the Netherlands

Correspondence: Norbert G. C. Huyer, Johanna Westerdijkplein 75, 2521 EN The Hague, The Netherlands. Tel: 31-628-373-202. E-mail: n.g.c.huyer@hhs.nl

Received: April 9, 2024 Accepted: June 15, 2024 Online Published: July 10, 2024

doi:10.5539/hes.v14n3p73 URL: https://doi.org/10.5539/hes.v14n3p73

Abstract

The advantages and drawbacks of components of flexible assessment have been studied mostly from the standpoint of students and, to a lesser extent, teachers. A gap persists in understanding the collective perspectives of teachers and students concerning flexible assessment. This study aimed to explore experiences and perspectives of students and teachers regarding flexible assessment within the specific context of nursing education. Seven focus groups comprised four sessions with teachers and three with students, each involving 5-8 participants. Results showed that students and teachers have a predominantly positive perspective towards flexible assessment. They acknowledge the opportunities that flexible assessment provides for diverse forms to present evidence. However, concerns were raised regarding the design of flexible assessments, issues of fairness in rating evidence, and the understanding among teachers and students regarding the assessment processes. Additionally, discussions focused on the perceived benefit of flexible assessments, particularly concerning the time investment required for their implementation and evaluation. In conclusion, the success of flexible assessments is contingent on the careful consideration of its design, ensuring equitable evaluation of evidence, and fostering comprehensive understanding among both teachers and students. Recognizing potential disparities in views of students and teachers offers valuable insights into the effectiveness of flexible assessment. Achieving a balance between the flexibility of assessment formats, aligned forms of evidence, and an appropriate rating methodology is crucial for effective implementation.

Keywords: assessment design, flexible education, forms of evidence, nursing education, quality criteria assessments, student experience

1. Introduction

Worldwide there is a growing focus on flexible learning programs (Leidl et al., 2020; Müller & Mildenberger, 2021; Swallow & Coates, 2004; Preston et al., 2010; Abdelaziz et al., 2011). Flexible learning programs better

match the needs of students who expect to work, learn, and study whenever and wherever they want (Wanner & Palmer, 2015). Therefore, these programs may be more attractive to a larger population of potential students (Wanner, et al., 2021). This is essential, because of shortages of professionals in several domains such as teaching and healthcare (Alutis, Bishaw & Frank, 2014; Ashiedu & Scott-Ladd, 2012; Sutcher, Darling-Hammond, & Carver-Thomas, 2019; Weinhold & Gurtner, 2014). In nursing, demographic changes and increases in care complexity have created a global shortage of nurses in general and of bachelor-educated nurses specifically (Dos Santos, 2020; Drennan & Ross, 2019; King et al., 2021; Marć et al., 2018; Vazquez-Calatayud et al., 2021). To expand the pool of registered bachelor nurses, flexible learning programs could be more inviting for nurses with an associate degree to continue their learning in a bachelor of nursing program (Leidl et al., 2020; Müller & Mildenberger, 2021) or for other professionals who consider nursing as a second career (Saitho et al., 2022). For them, flexible learning creates better opportunities to balance studies with work and family commitments (Kelly et al., 2009; Preston et al., 2010).

Flexible learning programs give students more responsibility in what, when, how and where they learn (Müller & Mildenberger, 2021; Snow Andrade & Alden-Rivers, 2019). However, there is also potential in flexible learning to make assessments more flexible. Ideally, students receive more responsibility in demonstrating what they have learned by providing flexible evidence of competency which can be assessed (Loony, 2009; Ryan & Tilbury, 2013; Wanner & Palmer, 2015). Through the option to submit various types of products, flexible assessment is more student-driven than traditional, less-flexible assessments (Wanner & Palmer, 2015; Irwin & Hepplestone, 2012). In this way, students can pursue their interests more freely, both in terms of content and form (Cowan, 2023). Flexible assessment could therefore positively influence students' engagement, motivation, grades, attitudes, and self-regulated learning skills (McGarry, et al., 2015; Pacharn et al., 2013). However, while the literature demonstrates that flexible assessment positively contributes to more student autonomy in their learning process (Snow Andrade & Alden-Rivers, 2019; Müller & Mildenberger, 2021), it can also lead to stress and anxiety due to the decision-making burden on students who may be unsure how best to choose their assessments (Cowan, 2023). This uncertainty of students is more pronounced by the fact that many students lack the skills in assessment necessary to make the appropriate decisions (Wanner & Palmer, 2015). Even though students welcome the chance to have more choice and control over their assessments, they still desire guidelines and frameworks to work within (Wanner & Palmer, 2015). These guidelines are particularly important if students past study experience consists of less flexible assessments (Irwin & Hepplestone, 2012). For teachers, flexible assessment involves an increased amount of extra work to manage an unpredictable multiplicity of assignments (Cowan, 2023) within limited planning time (Wanner et al., 2021).

While research has explored the advantages and drawbacks of components regarding flexible assessment, particularly for students and to a lesser extent for teachers, there remains a notable gap in understanding the perceptions and experiences of both teachers and students regarding flexible assessment. Previous research has primarily focused on flexible assessment within specific courses (Wanner & Palmer, 2015), or on specific outcomes such as student autonomy, motivation, and engagement (Edwards, 2020; Snow Andrade & Alden-Rivers, 2019), as well as the learning environment (Müller & Mildenberger, 2021). What is currently lacking is a comprehensive, broader study that investigates students' and teachers' experiences of flexible assessment within a full and flexible curriculum. Because flexible assessment is something that students and teachers primarily shape together, it is important to include both perceptions. Recognizing the potential disparities in students' and teachers' views on education (Murphy et al., 2011; Könings et al., 2014), pinpointing divergences and convergences in their perspectives can offer valuable insights into effective flexible assessment strategies. This led to the following main question; How do students and teachers experience flexible assessments in nursing education? The following sub-questions were addressed.

1. What benefits and challenges do students perceive in flexible assessments?
2. What benefits and challenges do teachers perceive in flexible assessments?
3. How do the perceptions of flexible assessments differ between students and teachers?
4. What improvements can be made to enhance the implementation of flexible assessments in nursing education?

This exploration is important for advancing the development and comprehension of flexible assessment programs, aiming to achieve an optimal balance between flexibility and the preservation of its inherent value.

2. Methods

A qualitative exploratory study design, using focus groups, was performed between October 2019 and June 2021.

Focus groups are a well-established method to explore perceptions (Hammarberg et al., 2016). We chose for focus group discussions not only to gather data on what participants themselves think, but also share and compare their experiences and views, enriching the data by providing information on the sources of similarities and differences (Morgan, 2019). A checklist for explicit and comprehensive reporting of qualitative research (Tong et al., 2007) was used to report all aspects of the study.

2.1 Participants

A purposeful sampling strategy was utilized. All 17 universities of applied sciences in the Netherlands having a flexible bachelor of nursing program that participated in the "National Experiment Learning Outcomes" (Government of The Netherlands, 2016; 2021) were contacted to participate in the study. Four universities of applied sciences responded positively and agreed to participate in this study. Thereafter, the participating universities recruited students and teachers. In each nursing program, both students and teachers were included in separate groups to prevent a power imbalance in the focus groups. This setup allowed students and teachers to openly discuss their experiences and perceptions without influencing each other (Morgan, 2019; Murphy et al., 2011; Könings et al., 2014). The focus groups had 5-8 people and were supervised by a moderator and an assistant moderator. Students differed in age, gender and previous education. In total seven focus groups were conducted, four with teachers and three with students (Table 1). Due to the workload during the COVID-19 period, one university failed to organize a student group.

Table 1. Overview of participants and universities of applied science

University	Position	Data collection Type	N	Age range	Form
A	Student group	Focus group 1	7	25-40	Live
	Teacher group	Focus group 2	8	30-55	Live
B	Student group	Focus group 3	5	25-35	Video conference
	Teacher group	Focus group 4	6	30-45	Video conference
C	Teacher group	Focus group 5	6	30-50	Video conference
D	Student group	Focus group 6	5	25-35	Video conference
	Teacher group	Focus group 7	5	30-45	Video conference

2.2 Data Collection

One focus group with teachers and one focus group with students were conducted live in a classroom at their university. The remaining focus groups were held by a video conference, due to the COVID-19 pandemic. Microsoft Teams was used to facilitate the video conferences. Video conferencing is an effective way to conduct focus groups (Flynn, 2018). Each focus group lasted between 60 and 80 minutes and all focus groups were audio, or, in case of online interviews, video recorded. Focus groups were conducted based on a protocol. At the start of every focus group participants were asked to independently rate the flexibility of assessments at their institution on a five point Likert-scale, where 1 stands for 'completely inflexible' and 5 stands for 'completely flexible'. Then results were discussed collectively. After this opening question an interview guide was followed to ensure the same topics were discussed in each focus group (see appendix 1 for interview guides). The two interview guides (one for students, one for teachers) were developed with the help of feedback from several teachers and students. The major topics were 1) participants' perceptions on the process of flexible assessments, 2) how different sorts of evidence were used and validated, and 3) how the quality of assessments were guaranteed. Teachers were specifically questioned about their current experiences in using flexible assessment and in assessing different forms of evidence of competence. The inquiry aimed to understand what sorts of evidence were provided and which evidence was deemed most important for grading. Students were particularly questioned about their experiences in providing various types of evidence of competence, the forms of evidence they used or would have preferred to use, how they selected their evidence, and how the quality of assessment was ensured.

All focus groups were moderated by the main researcher (NH). The second researcher (PR) checked whether all topics were adequately covered. Both researchers are experienced in focus group research. In the focus group with teachers of the moderators' own university, a third objective researcher from another university (VH), was asked to observe the moderator and assistant moderator, ensuring reliability and congruence of the collected data. No indications of subjectivity or bias due to familiarity with the interviewees were observed.

2.3 Data Analysis

Audio recorded focus group interviews were transcribed verbatim. Interview transcripts were analysed using

ATLAS.ti (ATLAS.ti, Version 9). Names of participants and universities were replaced by codes to ensure anonymity. A thematic analysis according to Braun and Clarke (2006; Byrne, 2022) consisting of six phases was used to analyse the data of the seven focus groups. First, the main researcher (NH) familiarised himself with the data by transcribing all the focus group material and reading the transcribed material. Then, data was individually processed by NH, JD and PR and meetings were held to discuss results and interpretation of data. The labelled content was grouped according to similarities and overlapping concepts of interest and then sorted into themes, describing the perception of both students and teachers regarding the effects of flexibility and quality of assessment.

2.4 Ethical Considerations

This research followed the Medical Research Involving Human Subjects Act of the Netherlands (Government of The Netherlands, 1998). This study concerns a non-medical non-intervention study, conducted with people who are able to express informed consent willingly and consciously. Therefore, no approval of a Medical Ethics committee is required and informed consent suffices.

Participation in each focus group was entirely voluntary with anonymous reporting. Before interviews were conducted, participants obtained e-mails with information explaining the content and aim of the study. Participants were informed that the focus group would be recorded and collected data would not be disclosed to any third party. All participants provided written or oral (video conference) informed consent. At the start of each focus group informed consent was readdressed and participants were reminded of their right to withdraw from the study at any time.

3. Results

The qualitative findings from the focus groups are presented as descriptive summaries and interpretations of themes, illustrated by representative, literal translated, quotes. Each quote is provided with a code that refers to the university of applied sciences (A, B, C or D), student group (SG) or teacher group (TG) and the focus group in question (see Table 1). An overview of the number of quotes per theme and focus group, accompanied by illustrative quotes, can be found in Appendix 2.

Theme 1 Design flexible assessment

Most students and teachers express a positive outlook on flexible assessment, recognizing the opportunities it provides for presenting evidence in diverse ways. Specifically, both students and teachers appreciate the flexibility in scheduling assessments:

> *"Because I think in our description of learning outcomes, which are the starting point for evidence, we leave room for very different interpretations of how students can demonstrate them. Stimulated by the different contexts in which they work, differing interpretations of learning outcomes emerge."* (B, TG, 3)

Students highlight that the option to submit evidence at regular intervals aligns well with their ability to plan flexibly:

> *"The flexibility of assessment is well organised, as of course in the flexible variant we are allowed to register for assessments every month."* (D, SG, 6)

Some teachers believe, however, that when given the opportunity to submit different types of evidence, students may find it challenging to choose between assessment formats:

> *"Within the "Quality of Care" module, we had the option that they could submit a film or a report, yet everyone submitted a report."* (D, TG,7)

Despite recognizing the potential benefits of flexible assessment, both students and teachers often raise concerns about its quality when they observe that assessment criteria do not align well with the evidence they provide. The consensus among most students and teachers is that assessment formats could be better aligned with learning outcomes and rubrics, especially considering the flexibility in the types of evidence and assessment formats that can be used. According to their perspective, a broader range of evidence and assessment formats should be eligible based on the learning outcomes:

> *"We are quite flexible when it comes to planning and when you can hand in assessments. But when I look at the aim of a flexible curriculum, where people could develop their own professional product based on learning outcomes and indicators, I see that it ended up being less flexible than we originally intended."* (B, TG, 4)

Many students find that the phrasing and description of learning outcomes play a decisive role in shaping the

assessment format, often restricting options to, for example, a written report. Students frequently encounter situations where learning outcomes and criteria/rubrics are either overly broad or too detailed. In instances of detailed rubrics, students feel constrained in their flexibility, as the criteria leave little room for alternative evidence. This often results in students adhering closely to the rubric, opting for the 'safest' and most obvious evidence to fulfil the criteria:

> *"And if you then look at the indicators and learning outcomes that are used at the university of applied sciences, there is a very simple one, for example; you have to be able to demonstrate that you can network, that you use your network. Yes you can interpret that in so many ways, and then you actually still have to guess."* *(B, SG, 3)*

> *"I have looked very much at the rubric, because I notice, that in the course, it is very much valued when assessing the evidence you hand in. It is ticked off exactly where the evidence does and does not comply."* (A, SG, 1)

Some teachers also say that the level of detail determines the flexibility of the assessments and say that criteria could be formulated more generally than is currently done. A desire is expressed to handle this differently:

> *"But, of course, the idea is always that you move towards rubrics that are not too detailed but also a bit more specific than the learning outcome'.* (D, TG,7)

In the design of rubrics/criteria teachers name that it is important to strive for clear and understandable criteria, however criteria are, according to most students, not always clear and understandable:

> *"We've talked a lot about the rubrics, and for me sometimes it's quite difficult when you read those rubrics, what do these mean? It's so specifically worded sometimes, and maybe that's very educational language, but it's quite difficult when you don't have a guideline to start with, to understand exactly what it's all about...?"* (A, SG,1)

Theme 2, Evidence and assessment

Most teachers find the use of diplomas as evidence 'reasonably easy', whereas they perceive work experience as evidence to be 'difficult'. In assessing evidence, some teachers employ criteria such as 'Variation', 'Relevance', 'Authenticity', 'Actuality' and 'Quantity' (VRAAQ-criteria) or with STARRT criteria (Situation, Task, Action, Result, Reflection and Theory, sometimes supplemented by Transfer). Typically, students are required to create a portfolio in consultation with their tutor, followed by a presentation and/or criterion-based interview during which additional questions are posed:

> *"The procedure is that I personally see if it fits the rubric and why I think it fits. That is assessed by 2 assessors and then you have to explain again in an oral interview why you think it fits, and then it is up to the assessors to decide whether it is indeed the case that you have reached the intended level.."* (A, SG,1)

Most students find it very difficult to find authentic evidence that matches the learning outcomes. In particular, students notice that teachers find it difficult to assess different evidence, like working experience. Students feel that their working experience is not sufficiently valued:

> *"Yes, that is very difficult, that is something you do in your daily life, but that is not easy to represent in documents or other evidence."* (B, SG, 3*)*

> *"...Lecturers are very reluctant to recognize that you are already functioning at bachelor level. I myself did two specializations for intensive care and emergency medicine."* (A, SG, 1*)*

Unlike the students, most teachers feel they appropriately recognize students' working experience, but there are also teachers who believe they should provide more substantial consideration to working experience:

> *"We also have nurses with an associate degree, some have been for a very, very long time. Who have been working in their oncology department for about 10 years, but can't get that valued in any way now".* (C, TG, 5)

Students encounter difficulties with teachers defining cognitive aspects at the bachelor level. Teachers often tend to overcompensate by requesting additional evidence, or conversely, not validating specific evidence to err on the side of caution. Teachers themselves acknowledge the challenge of defining determining criteria for cognitive aspects at the bachelor level. Conversely, teachers observe that students may too quickly assume that they are already performing at bachelor level when, in the teachers' opinion, this is not yet the case. Students express the ongoing challenge of presenting evidence distinctive of the bachelor level, as they perceive teachers to be uncertain in determining the appropriateness of their evidence. This uncertainty prompts teachers to request

additional evidence, leading students to rely on more traditional evidence and assessment formats:

> *"But also in the fields of clinical reasoning, people are very quick to assume; yes I have been a nurse for so long that I know all that. But that is precisely "the step up to the bachelor level" that students don't have. That is the crux that they have to learn to recognize that they are indeed nurses with an associate degree and that they are not at bachelor level yet."* (C, TG, 5)

Theme 3, Fairness

Teachers believe they are adequately trained and competent in evaluating different types of evidence. They assert that their competence ensures fair assessments, with decisions that they perceive as accurate and constant over time, maintaining uniformity across assessors and specific assessment situations. Assessors frequently engage in consultations where they discuss and calibrate their assessment. Conversely, students often feel that their evidence is incorrectly assessed as insufficient, and perceptions of assessment outcomes can vary over time. Some students acknowledge teachers' efforts but frequently experience grade discrepancies among different teachers, leading them to believe that grades are contingent on the individual teacher:

> *"I've seen it happen in class with texts being handed in and graded. One gets a 9 and the other a 5.5 and these are products that are almost exactly the same. I don't know exactly why. I think one just has a different background or a different way of reading or interpreting the criteria, but you can clearly see the difference."* (D, SG, 6)

Most teachers emphasize the importance of grading evidence consistently, applying the same criteria uniformly for all students. They invest effort in calibrating criteria with other teachers. However, at times, they face the challenge that students do not perceive gradings across various assessment formats as comparable. Some students confirm that there are indeed differences in the way that evidence is rated:

> *"Sometimes I hear students say; Yes, that criterium based interview is graded more holistically than a report. If you do not use a certain theory in a report, it will be rejected, but if you do not use this in an criterium based interview, it is sometimes rated as satisfactory."* (B, TG, 4)

Theme 4, Insight and Understanding

Insight and understanding are considered crucial because, as students gain the flexibility to choose from various evidence options, having the right information is essential for informed decision-making. Students need a clear understanding of the available choices to make appropriate selections. However, there is a divergence in perceptions between students and teachers regarding what constitutes the 'right' information and whether the information provided is sufficient. While teachers emphasize the importance of insight and understanding, they note that students may initially find it challenging to adapt to different assessment formats and may struggle with making diverse choices. Teachers themselves vary in their opinions about the adequacy of the information provided:

> *"I don't think the information we have really matches. That the provision of information to students may not be complete."* (D, SG, 6)

> *"Examples of different evidence and also for different settings and that we would then have a discussion about it. That would help enormously in making it clear to students where we aim to go."* (D, TG, 7)

Most students find that information about using different types of evidence is insufficient for making informed choices. They express this sentiment because they anticipated receiving more guidance in the process of selecting suitable evidence aligned with the learning outcomes. Students emphasize the importance of timely and sufficient information, incorporating good examples (best practices), and enhanced guidance by tutors. In offering solutions for improvement, students suggest ideas such as leveraging experienced students as 'student buddy' to provide additional support:

> *"But I think that it is precisely in the provision of information and the perception of what is expected that there is a lot to be gained and that can be done by students who have been through it before. I think it would also help if there was a very clear document saying what is expected of you".* (B, SG, 3)

Teachers acknowledge the significant role they play in guiding students' through the process of choosing different forms of evidence. However, this guidance is often expressed as a desirable goal, leaving questions about how to achieve it effectively. In considering this, some teachers recognize the potential for fostering students' self-regulatory abilities, a skill that some teachers feel is lacking in their students:

> *"Because in the beginning I actually expected students to ask themselves a learning question and to start looking for literature, but despite the fact that they are generally very enthusiastic, they very quickly look for*

the shortest route: what's in the rubric if there is one, and now I'm going to write. That does require a lot of coaching and critical questioning from us." (A, TG, 2)

Theme 5, added value

Students tend to value feedback on their evidence less positively when it comes solely from peers. They express a stronger preference for learning from expert feedback and formative assessments, provided the feedback is of high quality and contributes effectively to their learning for assessment:

"What I have seen is that students were allowed to give peer feedback, which I do think is quite subjective whether something is ok or not." (A, SG, 1)

The costs and efficiency of utilizing different types of evidence were frequently mentioned by both students and teachers. All teacher and student groups discussed the challenges of seeking exemptions for learning outcomes through the use of various types of evidence. Many students express that determining the right evidence to demonstrate their mastery of a learning outcome with diverse types of evidence is a labour-intensive process. They view this procedure as excessively bureaucratic and, in their opinion, does not outweigh the effort of attending regular lessons and undergoing 'standard' assessment. Teachers observe that students carefully weigh the effort involved in applying for an exemption against the potential benefits it may bring:

"All I can say about the evidence for the exemption is that I was hesitant to apply for an exemption. At one point the choice was whether to go for it or not. When I read that it would take between 80 and 100 hours to build and substantiate the evidence, I made the decision for myself: I'm not going to do it." (B, SG, 3)

Additionally, teachers note that when different types of evidence are utilized, teachers invest more time in assessments. This is partly because it is less predictable when, which type of evidence is chosen and, according to teachers, it also takes more time to evaluate, align and substantiate different forms of evidence:

"We are now in the assessment weeks, and because it is so flexible, you don't know in advance what kind of assessments students are going to take, so which assessments, how many assessments and which assessment format. So we don't know." (A, TG, 2)

4. Discussion

The aim of this study was to explore the experiences of students and teachers with flexible assessment in nursing education. Findings indicate that the majority of students and teachers hold positive views about flexible assessment, recognizing its potential of providing evidence for learning outcomes in diverse ways. However, it became evident that the principle of providing different kind of evidence is not yet established. Both students and teachers identify several areas for improvement in the design and implementation of flexible assessment.

The main similarities between students and teachers are that both believe that assessment formats could be better aligned with learning outcomes and rubrics, especially considering the flexibility in the types of evidence and assessment formats that can be used. Both groups find it difficult to define assessment criteria for cognitive aspects at the bachelor level. Additionally, both teachers and students face challenges in seeking exemptions for learning outcomes using various types of evidence. The main differences between students and teachers are that students, more than teachers, frequently express that learning outcomes and criteria/rubrics are either overly broad or too detailed. In cases of detailed rubrics, students often feel constrained in their flexibility. Another difference concerns the formulation of criteria: according to most students, the criteria are not always clear and understandable, while teachers believe they are transparent. Teachers also believe they are adequately trained and competent in evaluating different types of evidence, whereas students notice that teachers find it difficult to assess different evidence, often leading to discrepancies in grades among different teachers. The last main difference is the perception of what constitutes the 'right' information and whether the information provided and the guidance given are sufficient for making informed choices.

This study reveals that the success of flexible assessment is contingent upon the design of the assessment and the insight and understanding of teachers and students. Achieving an optimal balance between flexibility, alignment forms of evidence, and an appropriate assessment methodology proved to be conditional on these factors. These results align with findings of Baartman et al. (2006, 2007) who observed a similar effect in traditional (not flexible) assessment of competencies. Baartman et al. emphasize the importance of achieving a valid fitness for purpose as a foundational criterion before optimizing other assessment criteria. Fitness for purpose, stipulating that the forms of assessment used should align with learning outcomes, is a fundamental criterion that influences all other criteria (Baartman et al., 2006; 2007). In this study, a mismatch between expectations regarding the content of learning outcomes between teachers existed, which may implicate insufficient degrees of fitness for purpose. This contributed to teacher uncertainty about the bachelor level and the alignment of evidence. As a

result, teachers may seek reassurance or compensation, leading to asking for more additional evidence and the development of rubrics that are either too broad or too precise. The consequences of misaligned evidence combined with overly broad or precise rubrics could manifest in differences in comparability, a lack of fairness, a lack of transparency, inefficiency due to the extra time students spend using alternative forms of evidence, and an overall reduction in the intended flexibility of assessments.

The purpose and added value of a rubric, lies in providing transparency to the assessment (Panadero & Jonsson, 2013). However, students in this study report experiencing assessments and rubrics that are described in such intricate detail that they implicitly determine the format, limiting flexibility. On the contrary, when criteria are formulated overly broad, students may feel uncertain about their current level in their learning process and the development of necessary clinical competence (Finstad et al., 2022). Rubrics are often designed with a specific form of evidence in mind, leading to potential mismatches when items are scored that may not be present or are presented differently. This means that great demands are made on the teacher's cognitive flexibility to assess such completion and transfers. Formulating criteria that are appropriate for different types of evidence is an important topic that needs to be considered in flexible assessment, highlighting the need for more research in this area. Teachers acknowledge that overly detailed rubrics limit the degree of flexibility in choosing different types of evidence. Students tend to select the most obvious evidence to align with the rubrics and seek certainty, often subconsciously encouraged by teachers who use terms like ‘the student describes’. As a results, students feel torn between meeting the precise criteria of the learning outcome, which may 'already' align with a specific format, and their desire to make independent choices in assessment formats and interpretations, possibly leading to less engagement and satisfaction with their learning (Dumitru, 2021).

A correct balance between the degree of flexibility of assessment formats, aligned forms of evidence and an appropriate method of rating can stimulate students to choose from different assessment formats and experience a sense of increased involvement, ownership and responsibility for their learning (Edwards, 2020; Wanner & Palmer, 2015). This can potentially lead to positive impacts on motivation, grades, attitudes and, possibly, self-regulated learning skills (Pacharn et al., 2013). To achieve this, it is important to assure that multiple assessment formats (rubrics) are aligned with learning activities (Biggs, 2003, 2014). Additionally, careful consideration should be given to the degree of flexibility in assessment formats in time. Recognizing the importance of fitness for purpose, teachers should familiarize themselves with various formats and different types of evidence, considering their own preferences and perceptions shaped by historical assumptions about the effectiveness or appropriateness of assessment formats (Irwin & Hepplestone, 2012). Faculties also play a role by facilitating time and commitment to implement flexible and personalised assessment (Wanner et al., 2021). This contributes to positive teacher efficacy, influencing how teachers perceive their work and job satisfaction (Moe et al., 2010). Highly efficacious teachers have a positive influence on the students’ motivation and achievement (Mojavezi & Poodineh Tamiz, 2012).

When fitness for purpose is achieved, it becomes crucial to ensure clarity regarding learning outcomes, rubrics, different evidence/formats, and quality of the evidence to be utilized (Irwin & Hepplestone, 2012). This clarity is essential for students to gain insight and understanding of the choices they make in flexible assessment. In the context of flexible assessment, transparency is needed not only for summative assessments but also for formative and self-assessment, influencing academic performance positively (Boxham & Boyd, 2007; Baartman et al., 2006, 2007; Brown & Harris, 2013). The formative use of rubrics within flexible assessment improves student performance, such as increasing transparency, reducing anxiety, aiding the feedback process, improving student self-efficacy, or supporting student self-regulation (Panadero & Jonsson, 2013). Within flexible assessment, transparency takes on a nuanced role due to the variety of possible evidence. Students express a substantial need for clear examples of assessment forms accompanied by supporting evidence, explicitly outlining what is or isn’t sufficient for bachelor level. Building understanding and insight is emphasized through examples rather than the assessment frameworks. This can be achieved by providing a variety of examples of evidence that explicitly clarify when a learning outcome is at bachelor level/sufficient, creating a database of examples for the different types of learning outcomes without it becoming a fill-in-the-blank exercise. Transparency in providing students clear, accessible, and understandable benchmarks for developing and judging their work, will show students how their work is evaluated and what is expected of them (Ragupathi & Lee, 2020). Providing such examples could also help teachers in the process of accomplishing collective agreement on what is bachelor level and what is not. In this way, clear communication is fostered about assignments from the start and throughout the course, which proved very important (Wanner et al., 2021).

Several strengths and limitations of this study should be acknowledged. A strength of the study is that the findings are meaningful for a range of educational programs aiming for flexible assessment. Although the data

collection took place in nursing, and generalizability is not an aim of qualitative research in itself, the study informs other educational programs than nursing on students and teachers perspective on flexible assessment. Another strength is the variation in the degree of flexibility among the different participating universities of applied sciences, which may contribute to a more comprehensive discussion of certain themes. A limitation of this study is the participation of only 4 out of 17 universities in this study. The reluctance of more universities to participate could be attributed to the research taking place during the Covid-19 period. Because this study was conducted during the Covid-19 period, the majority of focus groups was held by means of a video meeting which, however, proved to be a good alternative (Flynn et al., 2018). Despite concerns, the assessment of data did not reveal any loss in quality. We believe that the video meetings did not negatively impact this focus group study (Flynn et al., 2018). Additionally, at one participating university, organizing a student focus group in this period was not feasible. However, this did not pose a problem because a sufficient degree of saturation was achieved in the other 3 student focus groups. A strength was that the participating universities were distributed across the country and shared the same basis of the national experiment learning outcomes (Government of The Netherlands, 2016, 2021), making the results transferable to other flexible education programs in the Netherlands. Finally, a second strength of the study is the composition of homogeneous focus groups, consisting of separate student and teacher groups. This ensured that teachers did not influence students and vice versa, which is a crucial aspect when exploring the differing perceptions of students and teachers (Murphy et al., 2011; Könings et al., 2014).

5. Implications for Practice

A primary consideration is to strive for a correct balance between the degree of flexibility of assessment formats, aligned evidence and an appropriate method of rating (Biggs, 2003, 2014). When this balance is achieved, students are motivated to select from a variety of evidence and assessment formats, resulting in an enhanced sense of involvement, ownership and responsibility for their learning. In the design of rubrics used in flexible assessment, the focus should be on determining the degree of flexibility for each learning outcome and defining the optimal balance in the degree of abstraction/precision of criteria to fit multiple types of evidence. This requires a methodical approach of thinking through and developing criteria. One potential strategy is to explore collaborative validation with teachers and students or engage in co-development with students (Fraile et al., 2016). Though this collaborative process, students and teachers can collectively generate a set of criteria, with students actively participating in a teacher-supported process of discussing disciplinary quality standards and priorities. Particularly for flexible assessment, the use of criteria in conjunction with exemplars can foster a shared understanding of criteria and how they might be applied (Tai et al., 2018).

The second implication pertains strengthening insight and understanding. This study reveals that due to the diverse range of evidence students can utilize, determining the assessment quality in advance becomes challenging, hindering teachers from effectively utilizing quality frameworks. The variety of possible evidence underscore that students have a great need for clear examples of assessments, with supporting evidence, in which it is clear which ones are or are not sufficient for which reasons. Transparency can be achieved by presenting a diverse range of possible evidence, explicitly indicating when a learning outcome is at the bachelor level and considered sufficient. Establishing a comprehensive database of examples for different types of learning outcomes and evidence is recommended. Examples serve as tangible instances of quality and offer students the chance to exercise their evaluative judgment (Tai et al., 2018). When collecting clear examples of flexible assessments, preferably in collaboration with students (Cockett & Jackson, 2018), it is crucial to align these examples with the learning outcomes and rubrics.

In addition to the database, a clear module guide can be developed, highlighting the learning outcomes and associated potential evidence. Alongside the use of a module guide and the sharing of best practices, participants in this study emphasized the essential role of tutors and experienced students as 'student buddies' to guide the process of making the right choices in flexible assessment and developing the necessary skills, a finding also suggested by Wanner & Palmer (2015) and shown to positively impact students' self-efficacy (Cilliers et al., 2010) in beliefs about their ability to accomplish specific tasks (Lunenburg, 2011). This approach ensures that examples contribute to the development of insight and understanding, rather than solely relying on assessment frameworks. Teachers should collaborate to agree on these examples, fostering clear communication about assignments from the start of the course and throughout. This method enhances self-regulated learning by incorporating formative assessment and self-assessment into the flexible assessment program (Granberg et al., 2021; Panadero et al., 2017, 2018). The implications outlined above enable students to better take control of their learning through self-regulation (Bandura, 1982), optimizing the advantages of flexible learning and flexible assessments.

A third implication related to insight and understanding focuses on enhancing teachers' competencies in flexible assessment, thereby bolstering teacher efficacy. Teachers should familiarize themselves with various forms of evidence and consider their own preferences and perceptions regarding learning outcomes, diverse evidence, rating, and aligned educational programs. Tutors can play a crucial role in guiding the decision-making process (Dumitru et al., 2022), using best practices from the database of examples for different types of learning outcomes and various evidence. For teachers, self-awareness of their self-efficacy (Moe et al., 2010; Mojavezi et al., 2012) is essential. Institutions should support teachers in developing their teacher efficacy by providing the necessary time and commitment to implement flexible assessment (Wanner & Palmer, 2015; Wanner et al., 2021). It is crucial that flexibility aligns with the instructor's capacity to manage an unpredictable amount of additional work (Cowan, 2023). This support is vital to ensure that teachers feel confident and competent in implementing flexible assessment practices.

6. Conclusions

This study revealed a predominantly positive outlook towards flexible assessment, indicating that a majority of students and teachers perceive flexible assessment as an opportunity to present evidence in diverse ways. The study highlights both similarities and differences between students and teachers regarding assessment practices. Both groups agree that assessment formats need better alignment with learning outcomes and rubrics, particularly given the flexibility in types of evidence and assessment methods. They also share difficulties in defining assessment criteria for cognitive aspects at the bachelor level and in navigating the process for seeking exemptions based on various evidence types. However, differences were observed: students more frequently criticize the breadth or detail of learning outcomes and rubrics, feeling restricted by overly detailed criteria. They also perceive assessment criteria as unclear, whereas teachers view them as transparent. Additionally, while teachers feel confident in their ability to evaluate diverse evidence, students observe inconsistencies in grading among different teachers and find the provided information and guidance insufficient for informed decision-making. It emerged that the effectiveness of flexible assessments is contingent upon achieving a nuanced balance in the design, as well as the insight and understanding of assessments by teachers and students. This balance is crucial for aligning the degree of flexibility in assessment formats with a diverse range of evidence and an appropriate rating methodology. This study contributes to the ongoing development and refinement of flexible assessment and enhancing the overall educational experience, not only in nursing education but also in broader educational contexts.

References

Abdelaziz, M., Samer Kamel, S., Karam, S., & Abdelrahman, A. (2011). Evaluation of E-learning program versus traditional lecture instruction for undergraduate nursing students in a faculty of nursing. *Teaching and Learning in Nursing, 6*, 50-58. https://doi.org/10.1016/j.teln.2010.10.003

Alutis, C., Bishaw, T., & Frank, M. W. (2014). The workforce for health in a globalized context - global shortages and international migration. *Global Health Action, 7*(2), 1-7. https://doi.org/10.3402/gha.v7.23611

Ashiedu, J. A., & Scott-Ladd, B. D. (2012). Understanding teacher attraction and retention drivers: Addressing teacher shortages. *Australian Journal of Teacher Education, 37*(11), 23-24. https://doi.org/10.14221/ajte.2012v37n11.1

Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Vleuten, C. P. M. van der (2006). The wheel of competency assessment: presenting quality criteria for competency assessment programs. *Studies in Educational Evaluation, 32*, 153-170. https://doi.org/10.1016/j.stueduc.2006.04.006

Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Vleuten, C. P. M. van der (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review, 2*, 114-129. https://doi.org/10.1016/j.edurev.2007.06.001

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147. https://doi.org/10.1037/0003-066X.37.2.122

Biggs, J. (2003). Aligning teaching for constructing learning. *Higher Education Academy, 1*(4), 1-4.

Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education, 1*, 5-22.

Bloxham, S., & Boyd, P. (2007). *Developing Effective Assessment in Higher Education: a practical guide*. Open University Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77-101. https://doi.org/10.1191/1478088706qp063oa

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). Thousand Oaks, CA: SAGE. https://doi.org/10.4135/9781452218649.n21

Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity, 56*, 1391-1412. https://doi.org/10.1007/s11135-021-01182-y

Cilliers, F. J., Schuwirth, L. W., Adendorff, H. J., Herman, N., & van der Vleuten, C. P. (2010). The mechanism of impact of summative assessment on medical students' learning. *Advances in Health Science Education, 15*(5), 695-715. https://dx.doi.org/10.1007%2Fs10459-010-9232-9

Cockett, A., & Jackson, C. (2018). The use of assessment rubrics to enhance feedback in higher education: An integrative literature review. *Nurse Education Today, 69*, 8-13. https://doi.org/10.1016/j.nedt.2018.06.022

Cowan, M. (2023). Flexible assessment; some benefits and costs for students and instructors. *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2023.2263668

Dos Santos, L. M. (2020). I Want to Become a Registered Nurse as a Non-Traditional, Returning, Evening, and Adult Student in a Community College: A Study of Career-Changing Nursing Students. *International Journal of Environmental Research and Public Health, 17*(16), 5652. https://doi.org/10.3390/ijerph17165652

Drennan, V. M., & Ross, F. (2019). Global nurse shortages-the facts, the impact and action for change. *British Medical Bulletin, 130*(1), 25-37. https://doi.org/10.1093/bmb/ldz014

Dumitru, T. C. (2021). Motivation and its challenges to learning. *Studia Universitatis Moldaviae, 9*(119), 43-48. https://doi.org/10.5281/zenodo.5779525

Dumitru Tăbăcaru, C.,Dumitru G. & Deaconu, M. B. (2022). Student's voice -What do students appreciate and expect to receive from a master program, a cross-sectional analysis, *Cogent Education, 9*(1), 2107298, https://doi.org/10.1080/2331186X.2022.2107298

Edwards, A. (2020). Playing To Their Strengths: Empowering Students With Flexible Assessment. *International Journal of Innovation in Science and Mathematics Education, 28*(4), 16-27. https://doi.org/10.30722/IJISME.28.04.002

Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: the effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation, 53*. 6-76. https://doi.org/10.1016/j.stueduc.2017.03.003

Finstad, I., Knutstad, U., Havnes, A., & Sagbakken, M. (2022). The paradox of an expected level: The assessment of nursing students during clinical practice - A qualitative study. *Nurse Education in Practice, 61*, 103332. https://doi.org/10.1016/j.nepr.2022.103332

Flynn, R., Albrecht, L., & Scott, S. D. (2018). Two Approaches to Focus Group Data Collection for Qualitative Health Research: Maximizing Resources and Data Quality. *International Journal of Qualitative Methods, 17*, 1-9. https://doi.org/10.1177%2F1609406917750781

Government of The Netherlands. (1998). Wet medisch-wetenschappelijk onderzoek met mensen. Retrieved from https://wetten.overheid.nl/BWBR0009408/2020-01-01

Government of The Netherlands (2016). *Handreiking pilots flexibilisering hoger onderwijs*. Den Haag Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from https://open.overheid.nl/documenten/ronl-archief-6ee6d669-b474-46b3-971f-71f78887f1de/pdf

Government of The Netherlands. (2021). *Overzicht van instellingen die deelnemen aan het experiment leeruitkomsten*. Retrieved from https://open.overheid.nl/documenten/ronl-f36efcd2-8edb-466c-aab5-ec1841fb960c/pdf

Granberg, C., Palm, T., & Palmberg, B. (2021). A case study of a formative assessment practice and the effects on students' self-regulated learning. *Studies in Educational Evaluation, 68*, 100955. https://doi.org/10.1016/j.stueduc.2020.100955

Hammarberg, K., Kirkman, M., & de Lacey, S., (2016). Qualitative research methods: when to use them and how to judge them. *Human Reproduction, 31*(3), 498-501. https://doi.org/10.1093/humrep/dev334

Irwin, B., & Hepplestone, S. (2012). Examining increased flexibility in assessment formats. *Assessment & Evaluation in Higher Education, 37*(7), 773- 785. https://doi.org/10.1080/02602938.2011.573842

Kelly, M., Lyng, C., McGrath, M., & Cannon, G. (2009). A multi-method study to determine the effectiveness of, and student attitudes to, online instructional videos for teaching clinical nursing skills. *Nurse Education Today, 29*(3), 292-300. https://doi.org/10.1016/j.nedt.2008.09.004

King, R., Taylor, B., Talpur, A., Jackson, C., Manley, K., Ashby, N., Tod, A., Ryan, T., Wood, E., Senek, M., & Robertson, S. (2021). Factors that optimise the impact of continuing professional development in nursing: A rapid evidence review. *Nurse Education Today, 98*, 1-15. https://doi.org/10.1016/j.nedt.2020.104652

Könings, K. D., Seidel, T., Brand-Gruwel, S., & van Merriënboer, J. J. (2014). Differences between students' and teachers' perceptions of education: Profiles to describe congruence and friction. *Instructional science, 42*(1), 11-30. https://doi.org/10.1007/s11251-013-9294-1

Leidl, D. M., Ritchie, L., & Moslemi, N. (2020). Blended learning in undergraduate nursing education - A scoping review. *Nurse Education Today, 86*. https://doi.org/10.1016/j.nedt.2019.104318

Looney, J. W. (2009). Assessment and Innovation in Education, *OECD Education Working Papers*, No. 24, OECD Publishing. https://doi.org/10.1787/222814543073

Lunenburg, F. C. (2011). Self-Efficacy in the Workplace: Implications for Motivation and Performance. *International Journal of management, Business and Administration, 14*(1), 1-6.

Marć, M., Bartosiewicz, A., Burzyńska, J., Chmiel, Z., & Januszewicz, P. (2018). A nursing shortage - a prospect of global and local policies. *International Nursing Review, 66*(1), 9-16. https://doi.org/10.1111/inr.12473

McGarry, B.J., Theobald, K., Lewis, P.A., & Coyer, F. (2015). Flexible learning design in curriculum delivery promotes student engagement and develops metacognitive learners: An integrated review. *Nurse Education Today, 35*(9), 66-73. https://doi.org/10.1016/j.nedt.2015.06.009

Moe, M., Pazzaglia, F., & Ronconi, L. (2010). When being able is not enough. The combined value of positive affect and self-efficacy for job satisfaction in teaching. *Teaching and Teacher Education, 26*(5), 1145-1153. https://doi.org/10.1016/j.tate.2010.02.010

Mojavezi, A., & Poodineh Tamiz, M. (2012). The Impact of Teacher Self-efficacy on the Students' Motivation and Achievement. *Theory and Practice in Language Studies, 2*(3). 483-491. https://doi.org/10.4304/tpls.2.3.483-491

Morgan, D. L. (2019). *Basic and Advanced focus groups*. London: Sage. https://doi.org/10.4135/9781071814307

Müller, C., & Mildenberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review, 34*. https://doi.org/10.1016/j.edurev.2021.100394

Murphy, L. M., Shelley, M. A., White, C. J., & Baumann, U. (2011). Tutor and student perceptions of what makes an effective distance language teacher. *Distance Education, 32*, 397-419. https://doi.org/10.1080/01587919.2011.610290

Pacharn, P., Bay, D., & Felton, S. (2013) The Impact of a Flexible Assessment System on Students' Motivation, Performance and Attitude. *Accounting Education, 22*(2), 147-167. https://doi.org/10.1080/09639284.2013.765292

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129-144. https://doi.org/10.1016/j.edurev.2013.01.002

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on selfregulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74-98. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Andrade, H., & Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *Australian Educational Researcher, 45*(1), 13-31. https://doi.org/10.1007/s13384-018-0258-y

Preston, G., Phillips, R., Gosper, M., McNeill, M., Woo, K., & Green, D. (2010). Web-based lecture technologies: Highlighting the changing nature of teaching and learning. *Australasian Journal of Educational Technology, 26*(6). https://doi.org/10.14742/ajet.1038

Ragupathi, K., & Lee, A. (2020). Beyond fairness and consistency in grading: The role of rubrics in higher education. *Diversity and inclusion in global higher education: Lessons from across Asia*, 73-95.

https://doi.org/10.1007/978-981-15-1628-3_3

Ryan, A., & Tilbury, D. (2013). *Flexible Pedagogies: new pedagogical ideas*. The Higher Education Academy, York.

Saitho, A., Shimoda, K., Kawabata, A., Oku, H., & Horiuchi, S. (2022). Evaluation of the first Accelerated Bachelor of Science in Nursing program as a second career in Japan. *Nurse Education Today, 111*. https://doi.org/10.1016/j.nedt.2022.105275

Snow Andrade, M., & Alden-Rivers, B. (2019). Developing a framework for sustainable growth of flexible learning opportunities, *Higher Education Pedagogies, 4*(1), 1-16. https://doi.org/10.1080/23752696.2018.1564879

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2019). Understanding teacher shortages: An analysis of teacher supply and demand in the United States. *Education Policy Analysis Archives, 27*(35). https://doi.org/10.14507/epaa.27.3696

Swallow, V., & Coates, M. (2004). Flexible education for new nursing roles: reflections on two approaches. *Nurse Education in Practice, 4*, 53-59. https://doi.org/10.1016/S1471-5953(03)00018-0

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education, 6*(3), 467-481. https://doi.org/10.1007/s10734-017-0220-3

Tong, A., Sainsbury, P., & Graig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349-357. https://doi.org/10.1093/intqhc/mzm042

Vazquez-Calatayud, M., Errasti-Ibarrondo, B., & Choperena, A. (2021). Nurses' continuing professional development: A systematic literature review. *Nurse Education Today, 50,* 1-9. https://doi.org/10.1016/j.nepr.2020.102963

Wanner, T., & Palmer, E. (2015). Personalising learning: Exploring student and teacher perceptions about flexible learning and assessment in a flipped university course. *Computer & Education, 88*, 354-369. https://doi.org/10.1016/j.compedu.2015.07.008

Wanner, T., Palmer, E., & Palmer, D. (2021). Flexible assessment and student empowerment: advantages and disadvantages - research from an Australian university. *Teaching in Higher Education*. https://doi.org/10.1080/13562517.2021.1989578

Weinhold, I., & Gurtner, S. (2014). Understanding shortages of sufficient health care in rural areas, *Health Policy, 118*(2), 201-214. https://doi.org/10.1016/j.healthpol.2014.07.018

Appendix

Appendix 1 Interview guides

Introduction

The same introduction is used in every focus group with students and teachers in which the moderator agrees up on the conditions and rules regarding the conditions of a good focus group discussion. For this, the substantive frameworks are explained and the rules of interaction are agreed.

The Moderator introduces himself and the assistant moderator. A introduction is given:

- The aim of the focus group is repeated; this study explores the perception of both students and teachers on the process of flexible assessments and quality validation.
- Explanation of the concepts flexible assessments/evidence and quality validation criteria is given
- The rules of the focus group are determined with each other with regard to safety, treating each other with respect and everyone should be able to contribute, so that each other is given the same opportunity to have input and space in time to respond.
- Duration and planning of the focus group is discussed

Interview guide teachers

Opening question

- 'Rate the flexibility of the assessments from 1 to 5, where 1 is not at all flexible and 5 is very flexible'.

- Why did you give this rating?

Main question 1

- what are your experiences in flexible assessment and receiving evidence from students to demonstrate (parts of) a learning outcome? And in what format do students submit evidence?

With the sub questions:

• Do you offer different kinds of assessment formats?

- What kind of evidence do you receive?

• What considerations do you make when assessing the evidence

• What quality requirements are set for the evidence?

• Which evidence do you think is the most important? Prioritize

• What are types of evidence that are not currently being used, but that you would like to validate?

• What are your experiences?

Main question 2

'How do you assess the evidence that students bring in to demonstrate that a learning outcome has been (partially) achieved'?

With the sub questions:

• Are specific assessment formats/procedures used to validate these?

• How do you assess the content of the evidence? Or what do you find important in this?

• What criteria/procedure do you use?

• What works well / less well / How would you like to do this better? Prioritize with teachers

Interview guide students

Opening question

- 'Rate the flexibility of the assessments from 1 to 5, where 1 is not at all flexible and 5 is very flexible'.
- Why did you give this rating?

Main question 1

''What are your experiences in flexible assessments and providing evidence to demonstrate (parts of) a learning outcome?

With the sub questions:

- Can you use different kind of assessment formats?
- Can you sufficiently use your own evidence?
- Which (type of) evidence do you use?
- What considerations do you make in choosing your evidence
- What quality requirements are set for the evidence?
- Which evidence do you think is the most important? Prioritize
- Are preconditions imposed on the (format) of the evidence?
- What is evidence that are not currently being used, but that you would like to have validated?
- What are your experiences?

Main question 2

''What does the procedure look like for submitting your evidence to demonstrate that you have already (partially) achieved a learning outcome?

With the sub questions:

- Are specific assessment formats/procedures used to validate these?
- What does it look like?

- How do you experience this procedure?
- What works well/less well/less here? (and how come)? P. Prioritize
- What are the consequences for further learning (formative)?
- Do you consider the effort for the process of flexible assessment in proportion to standard assessments?

Appendix 2; quotes themes

Themes	Focusgroup Students/teachers	University	Number of quotes	Example quotes
3.1 Design (think of: design learning outcome, rubric and flexibility)	Focus group 1 (SG) Focus group 2 (TG) Focus group 3 (SG) Focus group 4 (TG) Focus group 5 (TG) Focus group 6 (SG) Focus group 7 (TG)	A A B B C D D	17 18 11 23 20 3 26	*'But I think especially with those learning outcomes; they are just so difficult to interpret and determine for yourself whether your interpretation is correct'. (B, SG, 3)* 'W*e are increasingly able to make our assessments less detailed, which also creates more freedom for students to choose alternative assessment formats and to demonstrate what they are able of in a different way'. (D, TG, 7)* *'We initially said that everyone is completely free to submit as they want, that is still the case because the learning outcome and the indicators are leading, but then if you look in the rubric, it is often quite prescriptive; a student designates a specific method, so then it all becomes a bit more specific for the teacher and the student'. (B, TG, 4)* *'When I start, I have the rubric, I have the learning outcome, then I have learning activities. Then I think where do I start? Where should I go'? (A, SG, 1)* *'We talked a lot about the rubrics, and for me sometimes it's quite difficult when you read that rubric, what does it mean? It's so specifically named sometimes, and maybe that's very teaching language, but it's quite difficult when you don't have a guideline to start with, to understand; What is it about? What do you mean by this'? (A, SG,1)*
3.2 Evidence and Assessment (Think of: bachelor level, value previously acquired competence, form and content evidence, validating evidence, validating assessment)	Focus group 1 (SG) Focus group 2 (TG) Focus group 3 (SG) Focus group 4 (TG) Focus group 5 (TG) Focus group 6 (SG) Focus group 7 (TG)	A A B B C D D	47 67 27 52 35 45 62	*'Because you obviously have very good associate degree students and products that fall short at the bachelor level, and you're in a kind of grey area and I find that very difficult to say; so this is bachelor level'. (A, TG, 2)* *'What proves that you are really working at bachelor level? That's a very grey area and it's very difficult to determine the specific level.. (B, SG, 3)* *'I think you are already doing it and that is difficult if you for example already do a lot of networking in your daily work, then indeed it is difficult to prove that sometimes'. (B, SG, 3)* *'In the preparations, we both assessed the portfolio and we actually kind of rated it already and so asked additional questions, we also used the rubric to put next to it'. (A, TG, 2)* *'You will have to provide a substantial evidence. But how to do that remains complicated, doesn't it'? (C, TG, 5)* *' Student x already has much more work experience in healthcare, as a team leader and then you would think that should weigh more or maybe slightly more heavily, then some of my research and activities in a completely different area'. (B, SG, 3)*
3.3 Fairness (Think off: reliability and comparability)	Focus group 1 (SG) Focus group 2 (TG) Focus group 3 (SG) Focus group 4 (TG) Focus group 5 (TG) Focus group 6 (SG) Focus group 7	A A B B C D D	4 14 9 17 2 5 4	*'And I also find a big difference in how teachers assess that. That 1 teacher indicates during the lesson that what you are doing is already good and then you write it down and then you get it with a thick red line, it is insufficient so to speak and that I think, I really, really think that there is a lot to be gained within this university of applied sciences. Much more equal rating of teachers'. (B, SG, 3)* *'Well, that means a certain degree of objectivity, or intersubjectivity, that you come to the same judgment as different assessors. I think that is very important, that it is not arbitrary who your assessor is'. (C, TG, 5)* *'Some teachers are stricter in grading than others and so it's also a question of whether you meet the requirements, so that*

	(TG)			*also makes it tricky with assessment interviews'. (A, SG, 1)* *'We also get training if you want to know. We are trained assessors, which is also important to say. There is also quite a lot of practice, if you want you can practice quite a lot there, which is also good, at least I have experienced that as positive, instructive and especially also to be able to calibrate with each other'. (B, TG, 4)*
3.4 Insight and Understanding (Think of: information, guidance, using examples, expectations students and teachers)	Focus group 1 (SG) Focus group 2 (TG) Focus group 3 (SG) Focus group 4 (TG) Focus group 5 (TG) Focus group 6 (SG) Focus group 7 (TG)	A A B B C D D	32 17 24 6 3 15 27	*'It is precisely in that provision of information and the formation of an image of what is expected that, in my view, a lot of improvement can be made and that can be done by people who have been through it before, but a lot of information is also given and you can find that in a lot of places, I think it would also help if there was a very clear document of gosh this is what is expected of you'. (B, SG, 3)* *'So I think if you have good examples of what would be possible on the website, also to frame for us as teachers, we might find it easier to say: these are examples rated as sufficient'.* *(D, TG, 7)* *'But that we saw that students also very much want something to hold on to, so look for examples. So then you start giving examples and then you see that students still stick to those examples and that little flexibility is created'. (B. TG, 4)* *'I think that the study coach can have an important role in that. To see what evidence and experience the student has, to actually start that preliminary discussion. What does he or she have and what does he or she do in practice? And what does he or she think shows that level and how can you prove it? How can you show that'? (D, TG, 7)* *'Yes you do get some information on the intro day, but you get a lot of information there, so that was another part where I expected more support, or would have liked to have'. (B, SG, 3)*
3.5 Added value (Think of: self-regulation, feedback and Time & Costs)	Focus group 1 (SG) Focus group 2 (TG) Focus group 3 (SG) Focus group 4 (TG) Focus group 5 (TG) Focus group 6 (SG) Focus group 7 (TG)	A A B B C D D	23 21 13 9 5 9 10	*'Self-development; the tutor, who has a study group, should allow more formative assessments to take place so that the student knows "I am really ready to go into the assessment now'. (A, TG, 2)* *'That students have also seen the other's portfolio beforehand, so not just asking a question on the spot, but have thought of critical questions, because then it is also good preparation when they go to such an individual assessment. Or another assessment, but that they are trained almost like a teacher to look critically, ask critical questions to each other and that they can give feedback that gives onward feedback on what the student should still do'. (B, TG, 4)* *'I think that's an important step in this process to make it more accessible. For everyone. I think it's going to save a lot of time in considering: does or doesn't my evidence fit'. (B, SG, 3)*
				'Feedback to talk to each other, and that they discover for themselves; I am still missing something compared to the things described in the rubric'. (A, TG, 2)
				So we then do a kind of inventory in advance, but that's 3 weeks in advance and we've been so flexible that actually if something came up in between, that could be included. We said at a certain time on that day, whatever is handed in then will be assessed. And that's one thing that you don't know in advance, so you don't know how much work is coming at us and who is going to grade it. There is a tendency for students to hand in more and more, assessments have to be planned'.

				(A, TG, 2) *'Yes I actually worked full-time on the assessment in September and October, I actually put off everything else for me, already the first module. And that did make sure I got it right on paper, because that was a lot of work. I spent a lot of time on that. During that time, I had no time or space to devote to other subjects'. (B, SG, 3).* *'That many students then still choose to think then I'll just follow the classes and then I'll just take the 'regular assessment' because it ends up costing me as much time as requesting an alternative assessment'. (D, TG, 7)*