



Exploring Diagnostic Assessment as an Instrument for Learning Improvement

Mukhammad Bakhruddin^{a, 1*}; Sugito Muzaqi^{b, 2}; Kusaeri^{c, 3}; Lailah Isrofatun N.^{d, 4}; Mirwan A Taufiq^{e, 5}

^a Islamic Education Department, Faculty of Islamic Studies, Universitas Muhammadiyah Surabaya, Indonesia

^b Universitas Narotama Surabaya, Indonesia

^{c, e} UIN Sunan Ampel Surabaya, Indonesia

^d University of Exeter, England

¹ bakhruddin@um-surabaya.ac.id; ² Sugito.muzaqi@narotama.ac.id; ³ kusaeri@uinsby.ac.id; ⁴ Ln352@exeter.ac.uk; ⁵ mirwan@uinsby.ac.id

* Corresponding author

Article Info	Abstract
<p>Article History</p> <p>Received: March 19, 2023</p> <p>Revised: June 7, 2023</p> <p>Accepted: June 10, 2023</p> <p>Published: June 15, 2023</p> <p>Keyword: Diagnostic; assessment; learning</p> <p>Copyright (c) 2023 Mukhammad Bakhruddin, Sugito Muzaqi, Kusaeri, Lailah Isrofatun Nahdiah, & Mirwan Ahmad Taufiq</p> 	<p>This study aims to diagnose assessment as one of the learning improvement instruments. Assessment or assessment is a very urgent part of learning activities. From the diagnostic results, it is stated that the assessment has a positive impact on learning. One form of assessment that is often used is formative assessment. This research was conducted with a quantitative approach. Questionnaires, interviews, and documentation carried out data collection techniques. Questionnaires were distributed via Google form using a Likert scale. The research was conducted at madrasas within the scope of the East Java region. The number of respondents consisted of ninety students. The results showed that diagnostic assessment as a learning improvement instrument integrated with PAI was very effective in increasing the motivation of madrasa students. With an average of 4.21%, motivation shows its association with learning, 4.39% learning awareness, 4.31% learning tools, and 4.49% innovation, all related to Diagnostic assessment which also positively affects learning awareness. Diagnostic assessment also has a positive effect on learning awareness. The diagnostic assessment gives a positive response to learning tools or learning media. Diagnostic assessment has a positive influence on the Innovation of madrasa students.</p>
<p>How to Cite: Mukhammad Bakhruddin, Sugito Muzaqi, Kusaeri, Lailah Isrofatun Nahdiah, & Mirwan Ahmad Taufiq. (2023). Exploring Diagnostic Assesment as an Instrumen for Learning Improvement. <i>Progresiva : Jurnal Pemikiran Dan Pendidikan Islam</i>, 12(01), 53–68. https://doi.org/10.22219/progresiva.v12i01.25353</p>	

Introduction

Assessment is a component of the learning process (William, 2011). Because assessment will provide information related to ongoing learning (Struyven et al., 2005), educators will be able to reveal students' understandings, answers, comments, and so on. With such a holistic assessment, educators should be able to determine the learning progress of their students (Carrillo-de-la-Peña et al., 2009a) (Struyven et al., 2005) (Weurlander et al., 2012).

There are two types of assessments: formative and summative (Leenknecht et al., 2021). Two types of assessment are often used, namely formative and summative. Diagnostic assessment is another less critical evaluation type (He et al., 2022). Formative assessment is used to provide educators feedback to improve student learning achievement while learning is still ongoing. The summative assessment is completed after the lesson. Diagnostic assessments are given to determine the next level of education related to various things in preparation for learning, such as interests, learning styles, or student characteristics (I Kadek Mustika, 2022). Summative assessment is performed on student learning outcomes after attending lessons quarterly, one semester, or year-end. This paper will go over formative assessment in greater detail.

In practice, formative assessment becomes an informative tool for educators to discover their students' understanding so that educators can identify and improve/remediate their learning difficulties. Formative assessment can accurately estimate student learning, allowing you to track and report their progress (William, 2011). Formative assessment can take many forms, including written or oral, formal or informal. Educators, instructors, peers, and colleagues can conduct formative assessments. The formats, such as questions, conferences, interviews, dialogues, or discussions, can also differ.

Diagnostic evaluation is used to identify students' learning weaknesses so that appropriate treatment can be provided (Suharsimi Arikunto: 34). This assessment can determine the difficulties or obstacles students face during the learning process. This assessment is done for tutoring, remedial teaching, and finding other cases (Nana Sudjana 2010: 5).

Formative assessment can be designed to achieve various goals; however, it primarily generates feedback on student performance to improve learning (Weurlander et al., 2012). Feedback is an integral part of formative assessment, and students must understand and be able to act on it in order to improve their learning. On the other hand, students must be aware of the level or standard of performance required to improve their performance. They must also be able to compare the quality of their work to the required standard and devise strategies for closing gaps.

Diagnostic assessments assume that they improve student learning, supported by research findings. For example, they concluded diagnostic assessments that provide supportive student feedback improve their learning. It was recently demonstrated that students who participated in formative assessment performed better on the final exam than students who did not. Students from various fields of knowledge were involved in this research. The midterm exam serves as a formative assessment. Surprisingly, it appears that participation in formative assessment is more important than success, supporting the importance of feedback in formative assessment (Carrillo-de-la-Peña et al., 2009b). (Gilles et al., 2011; Yang & Embretson, 2007) argued for a model in which assessment is an essential

component of the teaching and learning process, and feedback plays a vital role in improving student achievement.

Several studies on diagnostic assessment include the research of Somaye Ketabi, Seyyed Mohammed Alavi, and Hamdollah Ravand with the title *Diagnostic Test Construction: Insights from Cognitive Diagnostic Modeling* (Ketabi, Somaye; Alavi, Seyyed Mohammed; Ravand, Hamdollah, 2021). In his study, it was explained that even though the Diagnostic Classification Model (DCM) was introduced to the education system several decades ago, it seems that this model was not used according to the original purpose for which it was designed. DCM uses mainly in analyzing large-scale non-diagnostic tests, and this model is rarely used in developing a Cognitive Diagnostic Assessment (CDA) from scratch. Despite the prevalence of "retrofit" CDA studies, "correct" CDA applications are believed to be rare because, first, a coherent framework for conducting such studies is not yet available, and second, researchers cannot analyze multiple DCMs for the same model according to index and criteria. This study summarizes the different types of DCM and reviews "true" and "retrofit" CDA studies. After examining the limitations of previous CDA studies, this study argues for the implications or application of Ravand and Baghaei's framework for conducting a "correct" CDA study. This framework is essential because it aligns with leading frameworks in educational assessment, such as the Cognitive Design System and the Assessment Triangle. It can also provide practical steps for test developers to conduct valid cognitive diagnostic tests.

The research conducted by Taufik Hidayat and Maemonah was entitled *Diagnostic Assessment: Analysis of Student Concentration Results in PAI Learning at SMP Plus Nusantara, Medan* (Taufik Hidayat, Maemonah, Juni-Desember). This study aims to see how the concentration of students participating in the learning process by paying attention to how much he accepts, rejects, and even students avoid the PAI learning process. The strategy carried out by educators is to carry out diagnostic assessment tests both in the cognitive, affective, and psychomotor domains. The diagnostic assessment examines problems that occur in students when the learning process is ongoing. The research was conducted using a descriptive qualitative method with the type of phenomenological research. The data collection technique in this study was through questionnaires by distributing Google forms to students at SMP Plus Medan City. Using a diagnostic test, the results showed that the concentration of students was almost as expected when viewed from the three aspects stated above, both from cognitive, affective, and psychomotor. This shows that students can do what educators hope for when participating in PAI learning.

According to Desmalia in his research *Implementation of Diagnostic Evaluation in Learning Islamic Religious Education (PAI) at State Vocational School 7 Rejang Lebong* (Desmalia, Desmalia, n.d.). This research uses diagnostic evaluation in learning Islamic Religious Education. This study aims to determine: learning difficulties in Islamic Religious Education learning, diagnostic Implementation in Islamic Religious Education learning, and handling of learning difficulties in Islamic Religious Education learning. This research was conducted at SMK Negeri 7 Lebong using qualitative research methods. The informants comprised school principals, Islamic Religious Education teachers, and several students: data collection techniques using interviews, observation, and documentation.

Moreover, analyzed using the theory of Miles and Huberman. The results of this study show that: First, known student learning difficulties include difficulties in reading the Qur'an and Hadith, difficulties in reading Arabic, memorizing short letters, difficulties

in practising PAI learning in material Aqidah Akhlak, Fiqh worship, and skiing. Second, a diagnostic evaluation is carried out based on students' learning difficulties, namely psychomotor, cognitive, and affective abilities. In addition, student learning difficulties are seen from internal and external factors and the results of student formative evaluations. Third, students with learning difficulties, such as not being able to read the Qur'an, will be guided by Islamic religious education teachers starting from learning to read Iqra' and those who cannot pray are also taught to pray appropriately, starting from movement to reciting prayers.

The urgency of diagnostic tests can also be studied in the research by Samsul Hadi, K. Ima Ismara, and Effendie Tanumihardja entitled Development of a Diagnostic Test System for Vocational Basic Competency Learning Difficulties for Vocational High School Students (Hadi et al., 2015). This study aims to: develop a diagnostic test system for learning difficulties for essential vocational competencies and determine the performance of a diagnostic test system for learning difficulties for essential vocational competencies for SMK students being developed. A diagnostic test system for learning difficulties for basic vocational competencies for vocational students has been developed with a client-server web architecture using the PHP, HTML, and Java programming languages with Apache, MySQL, and Chrome web browser supporting software. This system has three groups of users, namely admin, teachers, and students. The performance, design, and system adaptability analysis results have an average rating of 3.57 on a scale of 1-4, which is included in the perfect category. Therefore the system that has been developed can be used in further research, namely evaluation of system utilization.

Several studies on Islamic Religious Education assessment have been conducted, including one by Siti Lathifatus Sun'iyah (Sun'iyah, 2022). Because Islamic Religious Education cannot be separated from the application of calculating competence, this research explains how the Numerical Literacy Assessment in the minimum competency assessment (AKM) can be a reference for the Islamic Religious Education subject in developing its assessment. The content of the two literacy competencies represents cognitive abilities geared toward High Order Thinking Skills (HOTS). The Islamic Religious Education assessment incorporates AKM characteristics such as a computer-based test (CBT) and a more diverse set of questions. The National Assessment uses the character survey to assess students' social-emotional learning outcomes. The character values found in the Islamic Religious Education subject syllabus can be used to create this assessment instrument. This assessment is designed to be an assessment of learning. The findings of this research indicate that the AKM-oriented Islamic Religious Education assessment and Character Survey results are being followed up on to create learning that promotes critical thinking in solving societal issues.

Another research, entitled Implementation of the Assessment of Authentic al-Quran Hadist, was conducted (Matofiani & Prastowo, 2022). This was analyzed at authentic research on hadith maple al-Quran at the Islamic elementary school of al-Islam Giwangan in Yogyakarta. The research was conducted qualitatively. The findings demonstrated that authentic research for class XI students in the Al-Qur'an hadith subject at the Islamic Elementary School of Al-Islam Giwangan Yogyakarta was well implemented, including cognitive, affective, and psychomotor domains assessments. The teacher applied assignment techniques, quizzes, midterm tests, and final semester tests in the cognitive

domain. The teacher applied observation, self-assessment, and peer-assessment techniques specifically for the affective domain.

The psychomotor domain expresses opinions, discusses, and rearranges cut-off verses from hadith and projects. At the same time, the evaluation that needs to be improved is students' lack of enthusiasm for learning and inadequate classrooms. Nofrizal has also conducted other studies (Novrizal, 2022). The qualitative method was used in this research. According to the findings of this research, the national assessment has implications for the need for a focus on teaching Islamic Religious Education (Islamic Religious Education) in schools, which sharpens the linkages of the primary texts of Islamic teachings, the Qur'an and Hadith, to the contextual field. In addition, numeracy skills (mathematics) are required to study fiqh to solve inheritance cases.

Furthermore, literacy skills are critical for correctly understanding the text. Another implication of the national assessment for teaching Islamic Religious Education (Islamic Religious Education) and morals is how teaching can focus more on strengthening human character education in the Pancasila spirit. Agung Prihantono conducted another research entitled 'Failure of Authentic Assessment (Aa) in Islamic Religious Education Lessons in Elementary and Middle Schools' (Prihantono, 2021). This research used a systematic literature review approach, reviewing 27 articles with similar themes. The results are (1) failure of AA implementation in the form of at least the AA characteristics contained in the AA practices. (2) Because of a lack of socialization and AA training, the Islamic Religious Education teachers did not understand AA. (3) This failure has consequences for uneducated students' ability to apply knowledge, higher-order thinking skills, and attitudes in real-life situations, identify and solve complex Muslim problems, collaborate, and construct knowledge.

Muzlikhatun Umami conducted another research entitled 'Authentic Assessment of Islamic Religious Education and Moral Education Learning in the 2013 Curriculum' (Umami, 2018). According to the findings of this research, assessment in Islamic Religious Education is done in two ways: process assessment and learning outcome assessment. Assessment of the Learning Process employs an authentic approach that holistically evaluates student readiness, processes, and learning outcomes. The integrated assessment of the three components will describe students' capacity, style, and learning gain and produce an instructional effect (instructional effect) and a companion effect (nurturant effect) from learning.

Of the nine studies on diagnostic evaluation and assessment of Islamic Religious Education, none explicitly discussed assessment or assessment being an instrument in efforts to improve learning. Therefore, the research reviewed by this author is a differentiator from several other studies that intersect with the assessment of Islamic Religious Education.

Research Method

The method used is quantitative research, and data collection techniques are carried out utilizing observation, interviews, questionnaires, and document studies, supported by several applicable criteria in learning. Observation is used when data is taken regarding behaviour, natural phenomena, and the observed respondents are not too large. Observations are a reality or phenomena classified as relatively fixed, concrete, observable, and measurable. In making these observations, many characteristics must be considered, including 1) observations must be related to research objectives that have been set, 2) are systematic, 3) are quantitative, demand expertise and the results can later be tested with validity and reliability. The research process can be done in two ways: participating observation and non-participant observation. Observation involved researchers in daily activities, including learning activities for students to find patterns that already existed in the field so that data could be collected. At the same time, non-participant observation is carried out by not being involved in daily activities so that researchers are only independent observers.

In addition to observation, researchers also use interviews to collect data that is still minimal. There are two models of interviews: structured and unstructured. In a structured way, prepare materials before going to the field to obtain more accurate and structured data. The instruments built are more organized and more focused on the data collected so that it becomes valid and reliable. Meanwhile, in an unstructured way of getting information from the field according to the issues in the field, so that it is directly at the point of the problem or a complete variable, a representative party is needed in an object. Because the data obtained by researchers is not known with certainty, researchers must obtain a lot of information or answers from respondents. Questionnaire: this data collection is in the form of closed and open questions, sent via the postal service or the internet to go directly to the respondent. If the research scope is small, it can be delivered directly to the respondent.

Thus a good relationship and a conducive atmosphere have been established. Ultimately, the answers produced were accurate and valid for the respondents because they voluntarily provided answers that followed the researcher's questions. It is going to be. Low cost, ease of data collection, favourable statistical significance, low level of researcher subjectivity, and precise research findings (Werang, 2018). The Cronbach test method was used in this research for the normality test; this method is advantageous because it decides on the normality test when used on data with an average value less than the standard deviation (Leenknecht et al., 2021; Struyven et al., 2005). When deciding customarily distributed data, the Shafiro Wik test was used; if both t values, both skewness and kurtosis were in the value range $-1,96$ to $1,96$, the data were normally distributed; if both or only one of the skewness values or kurtosis is not within that range, the data were not normally distributed (Conderman et al., 2020; Tan Geok Shim et al., 2017).

The research instrument was a Diagnostic Assessment (Barker, 1988) using four indicators demonstrating the capacity to improve learning. The data generated by the measurement applied the commonly used Likert rating scale. The provided options such as Strongly Disagree (STS), Disagree (TS), Neutral (N), Agree (S), and Strongly Agree (SS). Number placement to represent opinions, from lowest to highest, such as STS: 1, TS: 2, N: 3, S: 4, and SS: 5. It is necessary to define the form of the statement, precisely the indicators that are collected, which include motivation, learning awareness, learning tools, and Innovation.

Result and Discussion

This research investigated the efficacy of diagnostic assessment in improving learning in Madrasah Aliyah students. However, instrument tests and descriptive analysis were performed prior to testing the effectiveness.

Instrument Evaluation

The validity and reliability of the questionnaire used as a data collection tool were first tested. This test assessed the questionnaire's suitability as a data collection tool. The following were the findings of the research questionnaire's validity and reliability tests:

Instrument Validity Test

Calculations were performed using correlation analysis to correlate each item with a total score. If the correlation coefficient value was more significant than r table = 0.2072, the indicator was valid for measuring the construct that was meant to be meaningful and was declared valid as a data collection tool. The validity test results are shown in the table below.

Table 1 Validity Test Results

Variable	Item Code	Correlation Coefficient	r table	Information
Motivation	X1.1	0,645	0,2072	Valid
	X1.2	0,744	0,2072	Valid
	X1.3	0,799	0,2072	Valid
	X1.4	0,798	0,2072	Valid
	X1.5	0,794	0,2072	Valid
Learning Awareness	X2.1	0,697	0,2072	Valid
	X2.2	0,767	0,2072	Valid
	X2.3	0,795	0,2072	Valid
	X2.4	0,741	0,2072	Valid
	X2.5	0,646	0,2072	Valid
learning tools	X3.1	0,759	0,2072	Valid
	X3.2	0,603	0,2072	Valid
	X3.3	0,738	0,2072	Valid
	X3.4	0,740	0,2072	Valid
	X3.5	0,472	0,2072	Valid
Innovation	X4.1	0,716	0,2072	Valid
	X4.2	0,653	0,2072	Valid
	X4.3	0,787	0,2072	Valid
	X4.4	0,795	0,2072	Valid

Based on the validity test results, it was discovered that all of the indicators in Table 1 produced a correlation coefficient value more significant than r table = 0.2072. As a result, all of the indicators in Table 1 have been declared valid and cannot be used as a data collection tool in this research.

Instrument Reliability Test

The Instrument Reliability Test was used to determine the instrument's consistency as a measuring tool so that a measurement can be trusted. Cronbach Alpha was used as a test. An instrument was more reliable if the alpha coefficient was more significant than 0.6. The table below summarizes the results of the questionnaire reliability test on all valid items based on the SPSS output.

Table 2 Reliability Test Results

Variable	Cronbach Alpha	Cut Off	Information
Motivation	0,704	0,600	Reliable
Learning Awareness	0,603	0,600	Reliable
learning tools	0,645	0,600	Reliable
Innovation	0,638	0,600	Reliable

According to the table, the Cronbach Alpha values for all variables in this research resulted in Cronbach Alpha values greater than 0.600, indicating that all question items in this research variable were stated to be consistent, reliable, and suitable for use as a data collection tool.

Respondents' Perception

Assessment categorization based on respondents' response scores. The assessment category was determined by the number of measurement scales used, ranging from one to five classifications.

$$P = \frac{X_{maks} - X_{min}}{b}$$

Where:

P = class length of each interval

X_{maks} = maximum value

X_{min} = minimum value

b = many classes

Table 3 shows the classification of assessment categories for the calculated average value based on the results of class length calculations for each interval.

Table 3 Classification of Rating Categories for Descriptive Statistics

Calculated Average Value	Rating Category
1 – 1,80	Very not good
1,81 – 2,60	Not good
2,61 – 3,20	Pretty good
3,21 – 4,20	Good
4,21 – 5,00	Very good

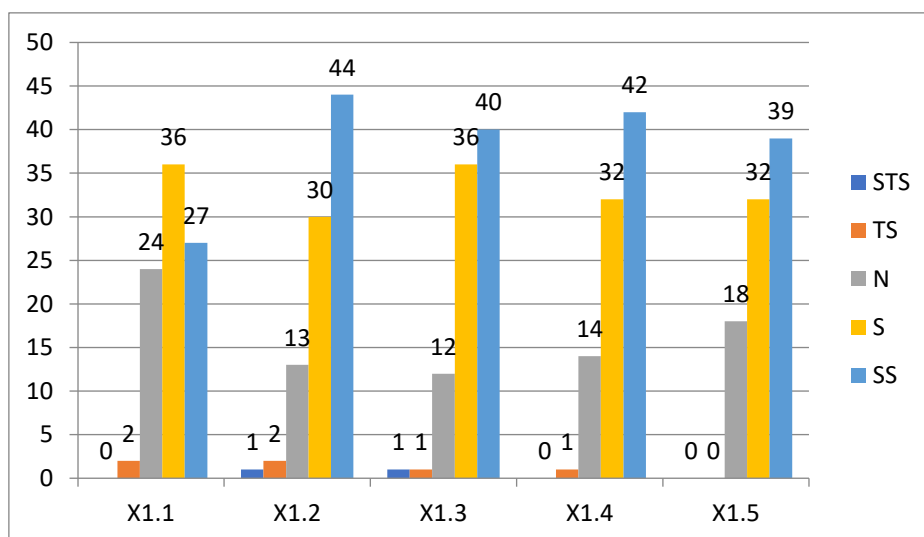
Source: *Processed Data Results*

Based on Table 3, the scale can be used as a guide to assess the outcomes of the existing questions, which were related to the existing variables and were discussed in this research. The following is a complete description of the respondents' perceptions of each variable:

Motivation Variable (X1)**Table 4 Respondents' Perceptions of Motivational Variables**

Item Code	STS		TS		N		S		SS		Average
	n	%	n	%	n	%	n	%	n	%	
I have high motivation to learning	0	0,0	2	2,2	24	26,7	36	40,0	27	30,0	3,99
I have good motivation for effective learning patterns	1	1,1	2	2,2	13	14,4	30	33,3	44	48,9	4,27
I am motivated to be an excellent example in the learning process	1	1,1	1	1,1	12	13,3	36	40,0	40	44,4	4,26
I am motivated by the challenge of a new learning model	0	0,0	1	1,1	14	15,6	32	35,6	42	46,7	4,29
I have a desire to succeed in acquiring knowledge	0	0,0	0	0,0	18	20,0	32	35,6	39	43,3	4,24
Average											4,21

According to the data above, the respondents' responses to the Motivation variable produce an average of 4.21. This means that respondents tended to rate the Diagnostic Assessment as an outstanding improvement in learning that was integrated with motivation. The question "I am motivated by the challenge of a new learning model" received the highest rating from respondents, with an average of 4.29. The question "I have high motivation in learning" received the lowest rating from respondents, with an average of 3.99. More information can be found in the graph below:



This graph shows that the acquisition of 0% states that they do not agree on the motivation variable, precisely on the item "I am highly motivated in learning, the acquisition at 24% lies in the high motivation item in learning, stating that it is neutral which does not take sides with others, while the statement agrees shows 36 % highest gain in the item "I have high motivation in learning." In this item, the gain is 27%, almost the same as

a neutral statement, so on the graph, the statement that agrees with the item I have high motivation in learning is essential.

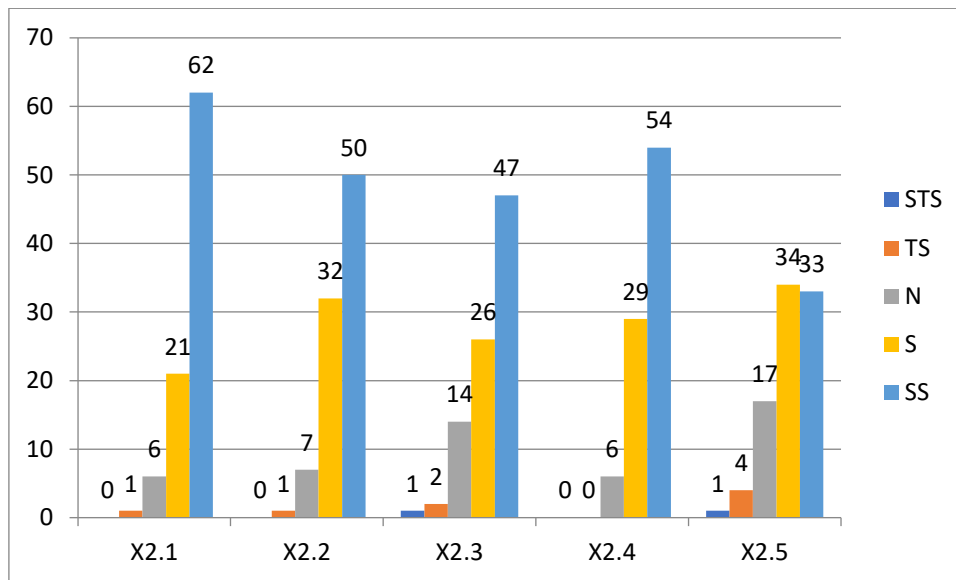
On the item "I have a good effective learning motivation pattern stating strongly disagree 1%, moderately disagree 2%, neutral 13%, agree 30%, and strongly agree 44%," this graph shows achievement in motivational statements in effective learning. In the item "I am motivated to be a good example in the learning process," the graph stated in statements that strongly disagree and disagreed 1%, those who chose neutral 12%, agreed 36% while strongly agreed 40%, it can be concluded motivational statements as examples the good ones on the graph are shown to have the highest response in society. In the item "I have the desire to succeed in acquiring knowledge," in this statement obtained strongly disagree and disagree 0%, those who chose neutral 18%, those who chose to agree 32% and those who chose strongly agree 39%, from some of the results of these statements it can be concluded that this motivational variable is closely related to effective learning.

Learning Awareness Variable (X2)

Table 5 Respondents' Perceptions of Learning Awareness Variables (X2)

Item Code	STS		TS		N		S		SS		Average
	n	%	N	%	N	%	n	%	n	%	
I like to study independently to increase awareness of learning	0	0,0	1	1,1	6	6,7	21	23,3	62	68,9	4,60
I was able to complete the tasks given by the teacher well	0	0,0	1	1,1	7	7,8	32	35,6	50	55,6	4,46
I am responsible for carrying out the tasks given by the teacher	1	1,1	2	2,2	14	15,6	26	28,9	47	52,2	4,29
I can reflect on the progress of science in the learning process	0	0,0	0	0,0	6	6,7	29	32,2	54	60,0	4,54
I can adapt to the learning process	1	1,1	4	4,4	17	18,9	34	37,8	33	36,7	4,06
Average											4,39

According to the data above, the respondent's responses to the Learning Awareness variable produced an average of 4.39. Respondents rated diagnostic assessments integrated with learning improvement and awareness as very good. The question "I like to learn independently to foster learning awareness" received the highest rating from respondents, with an average of 4.60. The question "I can adapt in the learning process." received the lowest rating from respondents, with an average of 4.06. More information can be found in the graph below:



This graph shows that the item "I like independent learning to raise awareness of learning obtained respondents with the following numbers: strongly disagree 0%, choose to disagree 1%, neutral 6%, agree 21%, then strongly agree 62%. It can be concluded from this statement that the results obtained from the respondents were very satisfactory, so the highest percentage was obtained, namely 62%. On the item "I was able to complete the task given by the teacher with good results for respondents who strongly disagreed 0%, disagree 1%, Neutral 7%, agreed 32%, and strongly agreed 50%, it can be concluded by selecting the highest respondent are very agree in this statement with indicated by 50% gain.

In the item "I am responsible for carrying out the tasks given by the teacher," in this statement, the respondents' choices were obtained with the following details: strongly disagree 1%, disagree 2%, neutral 14% and agree 26%, and strongly agree 47%. From the acquisition of the selection of respondents who have stated this, it can be concluded that the highest acquisition is the statement that strongly agrees with a percentage of 47%, so it can be concluded that this activity was achieved.

On the item "I can reflect on the progress of science in the learning process" in this statement, the respondents were: strongly disagree 0% disagree 0%, neutral 6%, agree 29% and strongly agree 54%. From the acquisition of these data, the highest achievement was obtained from strongly agreeing with the highest acquisition, namely 54%. It can be concluded that the acquisition was very satisfying.

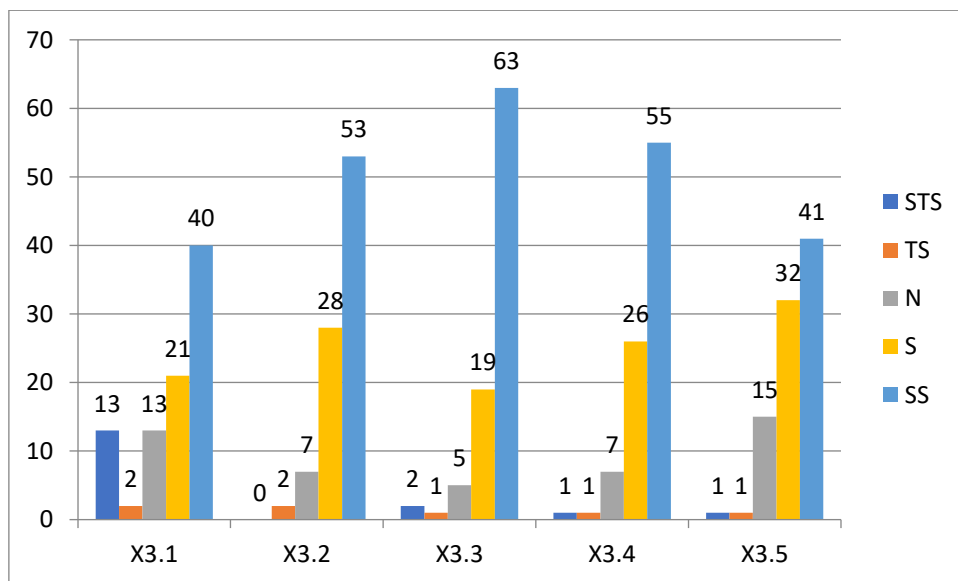
On the item "I can adapt in the learning process," this statement can be obtained strongly disagree 1%, disagree 4%, neutral 17% agree 34% and strongly agree 33%. This acquisition is the highest obtained by agreeing statements, so it can be concluded that the highest satisfaction is in the statements of respondents who agree, namely 34%

Tool for Learning Variable (X3)

Table 6 Respondents' Perceptions of Tools for Learning Variables (X3)

Item Code	STS		TS		N		TS		SS		Average
	n	%	n	%	N	%	n	%	n	%	
I can use media to accelerate learning	13	14,4	2	2,2	13	14,4	21	23,3	40	44,4	3,82
I am able to coexist with today's learning media	0	0,0	2	2,2	7	7,8	28	31,1	53	58,9	4,47
I am able to operate learning media properly and correctly	2	2,2	1	1,1	5	5,6	19	21,1	63	70,0	4,56
I can make tools around the classroom for learning media	1	1,1	1	1,1	7	7,8	26	28,9	55	61,1	4,48
I can formulate learning media for teaching and learning activities	1	1,1	1	1,1	15	16,7	32	35,6	41	45,6	4,23
Rata-Rata											4,31

According to the data above, the respondent's responses to the learning tool variable produced an average of 4.31. This means that respondents rated it very well after receiving a diagnostic assessment integrated with learning improvement tools. The question "I am able to operate learning media properly and correctly" received the highest respondent rating, with an average of 4.56. The question "I can use the media to accelerate learning" received the lowest rating from respondents, with an average of 3.82. More information can be found in the graph below:



The graph above it is explained with the average item in the statement "I am able to use media to accelerate learning," "I am able to coexist with today's learning media," "I am able

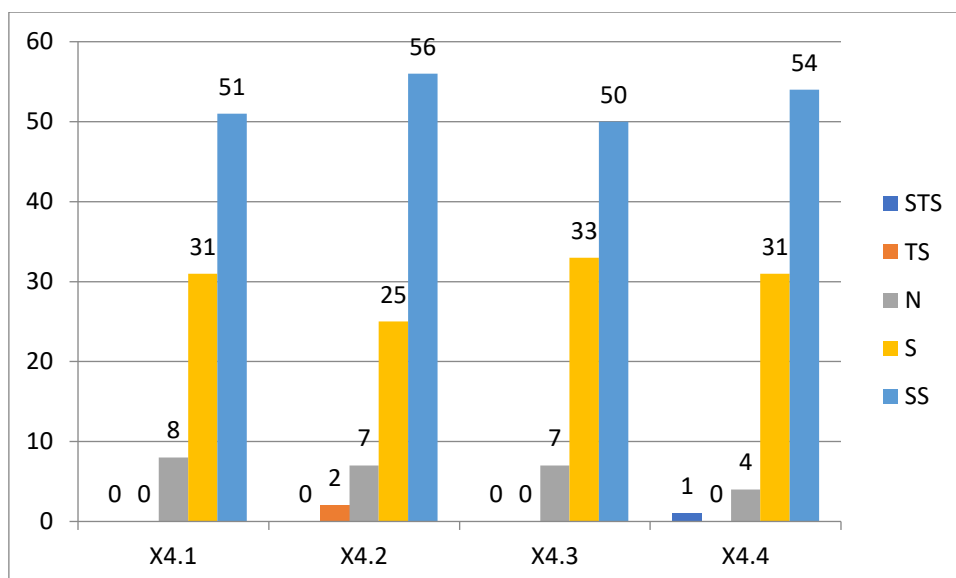
to operate learning media properly and correctly," "I am able to make tools around the classroom for learning media" "I am able to formulate learning media for teaching and learning activities" from several items the average has been achieved, namely 4.31%, it can be concluded that respondents were very satisfied with the learning measuring instrument at school.

Innovation Accommodation Variable (X4)

Table 7 Respondents' Perceptions of Innovation Variables (X4)

Item Code	STS		TS		N		S		SS		Average
	n	%	n	%	n	%	n	%	n	%	
I am able to provide creative work in the learning process	0	0,0	0	0,0	8	8,9	31	34,4	51	56,7	4,48
I can design an active learning model	0	0,0	2	2,2	7	7,8	25	27,8	56	62,2	4,50
I am active in creating creativity in teaching and learning efforts	0	0,0	0	0,0	7	7,8	33	36,7	50	55,6	4,48
I can use environmental materials for learning models	1	1,1	0	0,0	4	4,4	31	34,4	54	60,0	4,52
Average											4,49

According to the data above, respondents' responses to the Innovation variable resulted in an average of 4.49. This means that after receiving a diagnostic assessment that was integrated with learning improvement instruments that had Innovation in the excellent category, respondents tended to judge that. The question "I can use environmental materials for learning models" received the highest rating from respondents, with an average of 4.52. The question "I am active in creating creativity in teaching and learning efforts" received the lowest rating from respondents, with an average of 4.48. More information can be found in the graph below:



Pada grafik 7 bisa dijelaskan dengan responden inovasi bahwa rata rata keterlibatan dengan pembelajaran Pendidikan Agama Islam dapat dilihat dengan menggunakan prosentase "Saya mampu memberikan karya cipta dalam proses pembelajaran dengan skor

4,48 % dengan prosentase ini maka keefektifan terjadi, begitu juga pada “Saya mampu merancang model belajar yang aktif” rata rata respondennya 4,50 %, hal ini bisa dikatakan cukup sinergi, selanjutnya pada “Saya aktif dalam menciptakan kreatifitas dalam usaha belajar mengajar” maka hasil respondennya cukup baik skor rata ratanya 4,48 % di lanjutkan lagi dengan indicator pada “Saya mampu menggunakan bahan bahan lingkungan sekitar untuk model pembelajaran” dengan skor 4,52 %. Dari rata rata skor yang didapatnya maka sangat efektif sekali dalam pelaksanaan pembelajaran Pendidikan Agama Islam yang dikaitkan dengan pola diagnosis assessment sebagai alat ukur yang efektif.

Descriptive Analysis

The descriptive analysis described each research variable, including the minimum, maximum, median, average, and standard deviation values and the frequency distribution in the categorization results. The Descriptive Analysis results can be explained as follows:

Table 8
Research Variable Descriptive Analysis

Variable	Minimum	Maximum	Median	Mean	Std. Dev.
Motivation	2,20	5,00	4,2000	4,2106	0,62883
Learning Awareness	3,00	5,00	4,4000	4,3911	0,56720
Learning Tools	2,60	5,00	4,4000	4,3133	0,62821
Innovation	3,00	5,00	4,6250	4,4944	0,50137

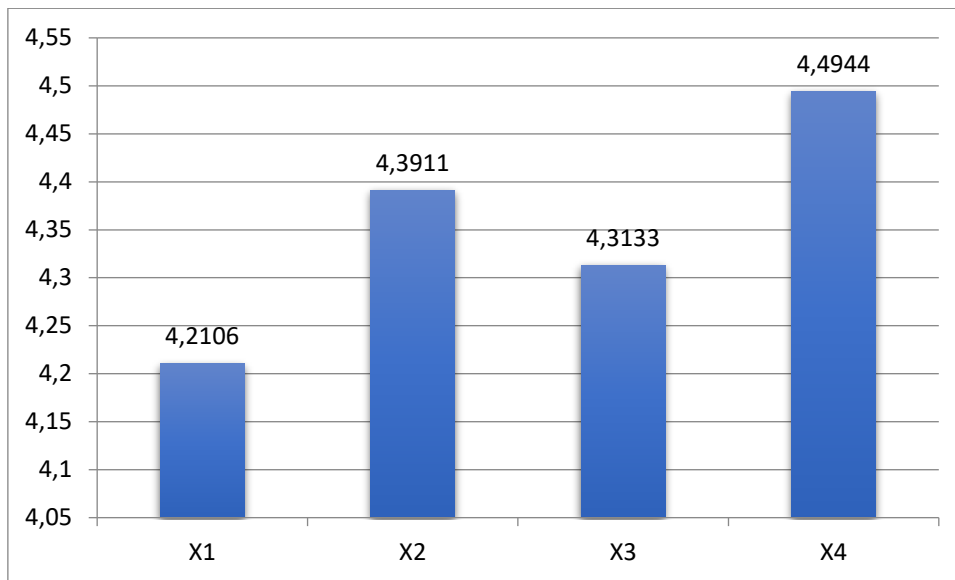
According to the data in Table 8, the lowest motivation was 2.20, and the highest was 5.00 out of 90 respondents. The respondents' average motivation score was 4.2106, the median was 4.2000, and the standard deviation was 0.62883. When the standard deviation was less than the average, it indicated that the variation in motivational values among respondents was slight.

Then, according to Table 1, the lowest learning awareness was 3.00, and the highest was 5.00, out of a total of 90 respondents. The average learning Awareness of the respondents was 4.3911, the median was 4.4000, and the standard deviation was 0.56720. The smaller standard deviation value than the average indicated that respondents' diversity of learning awareness values was small.

Furthermore, out of 90 respondents, the lowest was 2.60, and the highest was 5.00. The respondent's study result had an average Alaat score of 4.3133, a median of 4.4000, and a standard deviation of 0.62821. The fact that the standard deviation was less than the average indicated that the variation in the values of Tools for Learning among respondents was slight.

Moreover, it is known that the lowest Local Innovation was 2.20, and the highest was 5.00, based on a total of 90 respondents. The average value of respondents' innovations was 4.4944, the median was 4.6250, and the standard deviation was 0.50137. The fact that the standard deviation was less than the average indicated that the variation in the value of Innovation among respondents was slight.

The average variable graph produced the following results:



Normality Check

The normality assumption test determines whether or not variable data was normally distributed. The t-test assumes that the data was normally distributed. The Shapiro-Wilks test can be used to determine whether the data is normally distributed or not. The following is the hypothesis for testing the assumption of normalcy:

H_0 : Data is typically distributed

H_1 : Data is not normally distributed

According to the test criteria, the data was normally distributed if the probability value was greater than the significant alpha value of 5% or 0.05. The Shapiro-Wilks test yielded the following results when the normality assumption was tested:

Table 9 Normality Check

Variable	Statistic	Probability
Motivation	0,929	0,000
Learning Awareness	0,892	0,000
Learning Tools	0,895	0,000
Innovation	0,868	0,000

Source: Processed from SPSS results

The statistical probability of the Shapiro Wilks test was less than the significant alpha value of 5% or 0.05 based on the normality assumption test, so H_0 was accepted. This indicated that the data was not normally distributed. As a result, the assumption of normalcy was violated. Because the t-test cannot be used, alternative testing used the Wilcoxon analysis.

Analysis of the Effectiveness of Exploring Diagnostic Assessment as an Instrument for Improvement of Learning Associated with Effectiveness in Islamic Religious Education

Wilcoxon analysis was used in this research to assess the effectiveness of Diagnostic Assessment as an Instrument for Learning Improvement in Madrasah Aliyah students related to Islamic Religious Education. This was because the data was not normally distributed. If the probability value < level of significance ($\alpha = 5\%$) was obtained in the

Wilcoxon effectiveness test, it was stated to be significantly effective. The results of the relationship test with Wilcoxon analysis are shown below.

Table 10
Effectiveness Results with Wilcoxon Test

Variable	Statistic	Probability
Motivation	3,152	0,002
Learning Awareness	5,405	0,000
Learning Tools	4,296	0,000
Innovation	6,623	0,000

According to the analysis, the effectiveness test results obtained a significance value < level of significance ($\alpha = 5\%$ or 0.05) for all variables. As a result, it can be concluded that Diagnostic Assessment, as an instrument for improving learning integrated with Islamic Religious Education, was very effective in increasing motivation, learning awareness, learning tools, and Innovation in Madrasah students.

Conclusion

Assessment is a component of the learning process. Because assessment will be able to provide information related to ongoing learning. There are two types of assessments: formative and summative. Two types of assessment are often used, namely formative and summative. Another type of evaluation that is no less important is the diagnostic assessment. When diagnosing learning improvements, the author gives the visible treatment of students so that there are instruments that build those improvements. Motivation, awareness of independent learning, tools used to measure learning achievement, and Innovation are symptoms that appear in the diagnosis. This effort is made to improve good learning patterns, which can lead to good results. This is supported by several instruments that help strengthen the diagnosis and improve learning. The data that can be used as a reference in covering this is a statement supported by several indicators, namely motivation, with an average result of 4.21%, indicating a relationship between Diagnostic Assessment and motivation in improving learning. The second diagnostic indicator is learning awareness, with an average score of 4.39%, indicating that all Madrasah Aliyah students know independent learning.

Furthermore, with an average value of 4.31%, the third indicator is a learning tool. Significant improvements in learning can be achieved. The following four innovations have an average value of 4.49%. This demonstrates that Innovation is still relevant when combined with learning improvements due to participation in diagnostic assessments.

Suggestion

This research has several limitations, so further research should broaden its scope in terms of providing input in the future.

References

- Adachi, C., Tai, J. H.-M., & Dawson, P. (2018). Academics' perceptions of the benefits and challenges of self and peer assessment in higher education. *Assessment & Evaluation in Higher Education*, 43(2), 294–306. <https://doi.org/10.1080/02602938.2017.1339775>
- Alonzo, A. C. (2018). An argument for formative assessment with science learning progressions. *Applied Measurement in Education*, 31(2), 104–112. <https://doi.org/10.1080/08957347.2017.1408630>
- Barker, D. (1988). DIAGNOSTIC TESTS AND EXAMINATION RESULTS. *Assessment & Evaluation in Higher Education*, 13(1), 1–15. <https://doi.org/10.1080/0260293880130101>
- BOOK REVIEWS. (2006). *Assessment & Evaluation in Higher Education*, 31(4), 479–483. <https://doi.org/10.1080/02602930600679258>
- Burdick, A., & Willis, H. (2011). Digital learning, digital scholarship, and design thinking. *Interpreting Design Thinking*, 32(6), 546–556. <https://doi.org/10.1016/j.destud.2011.07.005>
- Carrillo-de-la-Peña, M. T., Baillès, E., Caseras, X., Martínez, À., Ortet, G., & Pérez, J. (2009a). Formative assessment and academic achievement in pre-graduate students of health sciences. *Advances in Health Sciences Education*, 14(1), 61–67. <https://doi.org/10.1007/s10459-007-9086-y>
- Carrillo-de-la-Peña, M. T., Baillès, E., Caseras, X., Martínez, À., Ortet, G., & Pérez, J. (2009b). Formative assessment and academic achievement in pre-graduate students of health sciences. *Advances in Health Sciences Education*, 14(1), 61–67. <https://doi.org/10.1007/s10459-007-9086-y>
- Conderman, G., Pinter, E., & Young, N. (2020). Formative Assessment Methods for Middle Level Classrooms. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 93(5), 233–240. <https://doi.org/10.1080/00098655.2020.1778615>
- Falchikov, N. (2013). *Improving Assessment through Student Involvement* (0 ed.). Routledge. <https://doi.org/10.4324/9780203220993>
- Gilles, J., Detroz, P., & Blais, J. (2011). An international online survey of the practices and perceptions of higher education professors with respect to the assessment of learning in the classroom. *Assessment & Evaluation in Higher Education*, 36(6), 719–733. <https://doi.org/10.1080/02602938.2010.484880>
- Huang, Y.-M., Huang, S.-H., & Wu, T.-T. (2014). Embedding diagnostic mechanisms in a digital game for learning mathematics. *Educational Technology Research and Development*, 62(2), 187–207. <https://doi.org/10.1007/s11423-013-9315-4>
- Husk, S. (2014). *The changeling soldier*.
- Kruit, P., Oostdam, R., van den Berg, E., & Schuitema, J. (2020a). Performance Assessment as a Diagnostic Tool for Science Teachers. *Research in Science Education*, 50(3), 1093–1117. <https://doi.org/10.1007/s11165-018-9724-9>

- Kruit, P., Oostdam, R., van den Berg, E., & Schuitema, J. (2020b). Performance Assessment as a Diagnostic Tool for Science Teachers. *Research in Science Education*, 50(3), 1093–1117. <https://doi.org/10.1007/s11165-018-9724-9>
- Leenknecht, M., Wijnia, L., Köhler, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: The role of students' motivation. *Assessment & Evaluation in Higher Education*, 46(2), 236–255. <https://doi.org/10.1080/02602938.2020.1765228>
- Regmi, K. D. (2017). Habermas, lifeworld and rationality: Towards a comprehensive model of lifelong learning. *International Journal of Lifelong Education*, 36(6), 679–695. <https://doi.org/10.1080/02601370.2017.1377776>
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review1. *Assessment & Evaluation in Higher Education*, 30(4), 325–341. <https://doi.org/10.1080/02602930500099102>
- Tan Geok Shim, G., Shakawi, A. M. H. A., & Azizan, F. L. (2017). Relationship between Students' Diagnostic Assessment and Achievement in a Pre-University Mathematics Course. *Journal of Education and Learning*, 6(4), 364. <https://doi.org/10.5539/jel.v6n4p364>
- Tang, F., & Zhan, P. (2021). Does Diagnostic Feedback Promote Learning? Evidence From a Longitudinal Cognitive Diagnostic Assessment. *AERA Open*, 7, 233285842110608. <https://doi.org/10.1177/23328584211060804>
- Voinea, L. (2018). FORMATIVE ASSESSMENT AS ASSESSMENT FOR LEARNING DEVELOPMENT. *Revista de Pedagogie - Journal of Pedagogy*, LXVI(1), 7–23. <https://doi.org/10.26755/RevPed/2018.1/7>
- Werang, B. R. (2018). Pengaruh Keterampilan Managerial Kepala Sekolah dan Iklim Sekolah Terhadap Komitmen Kerja Guru Sekolah Dasar Kristen di Kabupaten Boven Digoel. *Al Ibtida: Jurnal Pendidikan Guru MI*, 5(2), 159. <https://doi.org/10.24235/al.ibtida.snj.v5i2.2650>
- Weurlander, M., Söderberg, M., Scheja, M., Hult, H., & Wernerson, A. (2012). Exploring formative assessment as a tool for learning: Students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, 37(6), 747–760. <https://doi.org/10.1080/02602938.2011.572153>
- William, D. (2011). FORMATIVE ASSESSMENT: DEFINITIONS AND RELATIONSHIPS Dylan Wiliam, Institute of Education, University of London. *Assessment for Learning: Origins and Antecedents For*, 1–26.
- Yang, X., & Embretson, S. E. (2007). Construct Validity and Cognitive Diagnostic Assessment. In J. Leighton & M. Gierl (Eds.), *Cognitive Diagnostic Assessment for Education* (pp. 119–145). Cambridge University Press. <https://doi.org/10.1017/CBO9780511611186.005>
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45(4), 477–501. <https://doi.org/10.1023/A:1023967026413>