

<https://doi.org/10.17323/jle.2025.18317>

Navigating Motivation and Autonomy in Language Learning: Unveiling the Impact of Gradeless Assessment

Soroor Rostami ¹, Mansooreh Hosseinnia ² 

¹ Agder University, Kristiansand, Norway

² Azad University of Bojnourd, North Khorasan, Iran

ABSTRACT

Background: Traditional grading systems often undermine intrinsic motivation and learner autonomy in language education, prompting interest in alternatives like gradeless assessment. While descriptive feedback and self-correction show promise in enhancing learner development, their impact on specific motivational constructs, such as ideal and ought-to L2 selves, remains underexplored, particularly in non-Western EFL contexts. This study addresses this gap by examining how gradeless assessment influences motivation and autonomy, contributing to global debates on formative assessment practices.

Purpose: This study examines the effects of a gradeless assessment approach, utilizing descriptive feedback and self-correction, compared to conventional grading, on the motivation (ideal and ought-to L2 selves) and autonomy of EFL learners. It aims to address the pedagogical challenge of designing assessments that foster sustained motivation and learner agency, particularly in diverse educational settings like Iran, to inform global language education practices.

Method: Forty-one upper-intermediate EFL students, selected via the Quick Placement Test, were randomly assigned to an experimental (gradeless, $n=21$) or control (graded, $n=20$) group for a reading and writing course. Both groups received identical instruction and materials from the same instructor, differing only in assessment: the experimental group received written descriptive feedback with self-correction opportunities, while the control group received numerical grades. Pre- and post-test questionnaires assessed motivation and autonomy, with data analyzed using independent-sample t-tests.

Results: The gradeless approach significantly enhanced ideal L2 self-motivation ($t = -2.70$, $p < .05$, $\eta^2 = .15$, 9.8% mean increase) and autonomy ($t = -2.06$, $p < .05$, $\eta^2 = .15$, 1.1% mean increase), but reduced ought-to L2 self-motivation ($t = 3.05$, $p < .01$, $\eta^2 = 6.2\%$ mean decrease), suggesting a shift toward ideal L2 self-motivation.

Conclusion: These findings highlight the potential of gradeless, feedback-oriented assessment to foster intrinsic motivation and autonomy in Iranian EFL contexts, suggesting a need to reconsider traditional grading in favor of formative strategies. However, due to the study's cultural and contextual specificity, generalizability should be approached cautiously, warranting further research across diverse settings.

KEYWORDS

gradeless assessment; descriptive feedback; self-correction; L2 motivation, autonomy

Citation: Rostami, S., & Hosseinnia, M. (2025). Navigating motivation and autonomy in language learning: unveiling the impact of gradeless assessment. *Journal of Language and Education*, 11(3), 127-137. <https://doi.org/10.17323/jle.2025.18317>

Correspondence:
Mansooreh Hosseinnia,
mansoorehhosseinnia@yahoo.com

Received: November 10, 2023

Accepted: September 18, 2025

Published: September 30, 2025

INTRODUCTION

Traditional grading systems, which emphasize numerical scores, have been widely critiqued for their psychological and pedagogical limitations. The grading system remains a focal point of controversy and robust debate (Cain, Medina, Romanelli, & Persky, 2022; Jaschik,

2009; Lim, 2024; McMorran & Ragupathi, 2020). Research indicates that grades have detrimental psychological effects (e.g., Goos, Pipa, & Peixoto, 2021; Linder, Gerdtham, & Heckley, 2023; Marquez & Garcia, 2023; Supiano, 2019), raising concerns about the credibility of the conventional "A" to "F" grading system, highlighting well-documented trends in



grade inflation patterns (Rojstaczer & Healy, 2012). Jaschik (2009) encapsulated this debate by arguing that grades often create a superficial semblance of coherence. These findings have prompted growing interest in alternative assessment approaches, such as gradeless assessment, which prioritizes descriptive feedback and self-correction to foster learner-centered environments. In English as a Foreign Language (EFL) contexts, where motivation and autonomy are critical for sustained language acquisition, such alternatives hold particular promise for addressing the limitations of conventional grading (Dörnyei, 2005; Lamb, 2010).

The detrimental effects of numerical grading have been well-documented. Butler (1988) found that grades, even when accompanied by comments, can negatively impact students' engagement and performance by fostering a performance-oriented mindset. Similarly, Pulfrey, Buchs, and Butera (2011) argue that grades often operate through external pressures, such as fear of failure or peer competition, which can undermine students' willingness to engage deeply with learning material. In contrast, descriptive feedback, which provides specific guidance on how to improve without assigning scores, has been shown to enhance student success by focusing on learning processes (Brookhart, 2008; Belton, 2022). Studies by Guskey and Lee (2013) and Sackstein (2015) further challenge the necessity of grades, suggesting that formative feedback can support learning without numerical evaluations. Despite these insights, the application of gradeless assessment in diverse EFL contexts remains underexplored, particularly in terms of its impact on specific psychological constructs like motivation and autonomy.

Motivation and autonomy are pivotal in language learning, shaping students' academic success and long-term engagement (Lerner et al., 2022; Pratomo, & Kuswati, 2022; Quesada et al., 2025; Rostami et al., 2015; Smith et al., 2023). It seems that these two psychological variables are influenced by the grading system (Mendoza, Yan, & King, 2023; Muho & Taraj, 2022; Wu et al., 2014; Zedan, 2021). Dörnyei's (2005) L2 Motivational Self System (L2MSS) provides a robust framework for understanding motivation, conceptualizing it through the ideal L2 self (learners' internalized aspirations as language users) and the ought-to L2 self (external obligations driving language learning). Similarly, Kluger and DeNisi's (1996) Feedback Intervention Theory (FIT) posits that feedback focused on task improvement, rather than performance evaluation, enhances learners' engagement and self-regulation. These frameworks suggest that assessment practices can significantly influence learners' motivational orientations and autonomous behaviors. For instance, descriptive feedback aligned with learners' goals may strengthen their ideal L2 self, while reducing reliance on external pressures associated with the ought-to L2 self. Likewise, autonomy-supportive assessments, which en-

courage self-monitoring and reflection, can foster learners' sense of control and responsibility (Lamb, 2010).

Despite the growing body of research on formative assessment, a critical gap remains in understanding its impact on L2 motivation and autonomy in non-Western EFL contexts. Most studies have focused on Western educational settings, where cultural and institutional norms differ significantly from collectivist societies like Iran, where grading systems dominate instructional practices (Tavallali & Marzban, 2015). In Iranian EFL classrooms, traditional grading often limits opportunities for learner agency, as students are conditioned to prioritize scores over learning processes (Rostami et al., 2015). Moreover, while L2MSS and FIT have been widely applied to study motivation and feedback, their integration in examining gradeless assessment in non-Western contexts remains limited. This gap is particularly significant given the unique cultural and educational challenges in Iran, where collectivist values and exam-driven curricula may influence how learners respond to alternative assessment practices.

This study addresses this gap by investigating the effects of a gradeless assessment approach, characterized by descriptive feedback and self-correction opportunities, compared to conventional grading, on the motivation (ideal and ought-to L2 selves) and autonomy of Iranian EFL learners. Drawing on L2MSS, it predicts that gradeless assessment will enhance ideal L2 self-motivation by aligning feedback with learners' internalized goals, potentially reducing ought-to L2 self-motivation by minimizing external pressures. FIT further suggests that descriptive feedback will promote autonomy by directing learners' attention to task improvement and self-regulation rather than performance evaluation. By exploring these constructs in an underexplored non-Western context, this study contributes to global discussions on assessment reform, L2 motivation, and autonomy-supportive pedagogy. It aims to provide empirical evidence to inform the design of assessment practices that foster sustained motivation and learner agency, offering implications for language education policies and teacher training worldwide.

LITERATURE REVIEW

Psychological Impacts of Traditional and Gradeless Assessment

Traditional grading systems in education have faced significant criticism for their psychological and pedagogical limitations. Scholars argue that numerical grades increase student anxiety, discourage risk-taking, and shift focus from learning to performance outcomes (Butler, 1988; Pulfrey et al., 2011). For instance, Butler (1988) found that grades, even when accompanied by comments, foster a perfor-

mance-oriented mindset, reducing engagement. Conversely, gradeless assessment, emphasizing descriptive feedback and self-correction, has shown promise in mitigating these issues. Sackstein (2015) reported that gradeless classrooms empowered lower-achieving students by reducing stigma, though high-achievers struggled to adapt. Similarly, Benz¹ (2019) found that gradeless assessment reduced anxiety and reliance on external rewards, fostering a learning-oriented environment. However, these studies, often qualitative and Western-centric, lack experimental rigor to establish causal links between gradeless assessment and psychological outcomes like motivation and autonomy, highlighting the need for more robust investigations.

Feedback and L2 Motivational Self System

Dörnyei's (2005) L2MSS provides a framework for understanding how assessment practices influence language learning motivation through the ideal L2 self (learners' internalized aspirations) and ought-to L2 self (external obligations). Research suggests that descriptive feedback aligns with the ideal L2 self by supporting learners' intrinsic goals, while grades reinforce the ought-to L2 self through external pressures (Butler & Nisan, 1986; Wu et al., 2014). For example, Butler and Nisan (1986) found that task-related feedback enhanced engagement and outperformed grade-based systems. Wu et al. (2014) further noted that descriptive feedback fosters a task-involving environment, unlike grades, which promote competition and anxiety. However, these studies rarely apply L2MSS to EFL contexts or examine how feedback types differentially impact ideal and ought-to L2 selves, leaving a gap in understanding their motivational dynamics in language learning.

Assessment and Autonomy in EFL Contexts

Learner autonomy, defined as the capacity to take responsibility for one's learning (Dickinson, 1987), is closely tied to assessment practices. Lamb (2010) distinguishes between "assessment of autonomy" (measuring autonomy) and "assessment for autonomy" (promoting it through self-monitoring and reflection). Descriptive feedback and self-correction, as autonomy-supportive strategies, encourage learners to engage actively in their learning process (Pishghadam et al., 2011). In the Iranian EFL context, where grading systems dominate, studies like Tavallali and Marzban (2015) highlight limited autonomy due to exam-driven curricula. Razavi (2016) found that descriptive evaluation in Iranian primary schools faced challenges, such as inadequate teacher training and weak student self-evaluation skills. These studies,

while insightful, are often small-scale and qualitative, limiting their ability to generalize or address autonomy in relation to specific motivational frameworks like L2MSS in non-Western settings.

Methodological Gaps and the Need for Experimental Evidence

Despite the growing interest in gradeless assessment, significant methodological and contextual gaps persist. Most studies, such as Sackstein (2015, 2018²) and Spencer (2017)³, rely on qualitative or observational designs, which, while rich in detail, cannot establish causal relationships between assessment practices and psychological outcomes. Moreover, their focus on Western contexts overlooks cultural nuances in non-Western EFL settings like Iran, where collectivist values and rigid grading systems shape learner experiences (Mortazavizadeh et al., 2017). No experimental study to date has examined the impact of gradeless assessment on learners' ideal and ought-to L2 selves and autonomy in non-Western EFL contexts using a theory-driven design. This study addresses these gaps by employing an experimental approach, grounded in L2MSS and FIT (Kluger & DeNisi, 1996), to investigate the effects of gradeless assessment (descriptive feedback and self-correction) versus traditional grading on Iranian EFL learners' motivation and autonomy. It poses the following research questions:

RQ1: Does gradeless assessment (using descriptive feedback) have a statistically significant impact on Iranian EFL learners' ideal and ought-to L2 selves?

RQ2: Does gradeless assessment (using descriptive feedback) have a statistically significant impact on Iranian EFL learners' autonomy?

METHOD

Participants

To achieve the objectives of this study, 47 female students, aged between 21 and 33 ($M=25.47$, $SD=3.19$), were purposively selected from upper-intermediate level classes at a private language institute in Mashhad, Iran. The non-random selection aimed to ensure practical considerations. Six participants were excluded after the placement test, resulting in 41 homogenized individuals who were randomly divided into an experimental group (21 participants) and a control group (20 participants). The selection of up-

¹ Benz, G. (2019, October 31). *Going gradeless: A liberation from anxiety*. Retrieved from <https://www.teachersgoinggradeless.com/blog/anxiety-benz>

² Sackstein, S. (2018). Earning good grades versus learning. https://blogs.edweek.org/teachers/work_in_progress/2018/12/earning_good_grades_versus_learning.html

³ Spencer, K. (2017). A new kind of classroom: No grades, no failing, no hurry. *The New York Times*. <https://www.nytimes.com/2017/08/11/nyregion/mastery-based-learning-no-grades.html>

per-intermediate students was informed by both practical and pedagogical considerations. Upper-intermediate learners possess sufficient linguistic competence to comprehend and respond to complex questionnaire items, ensuring reliable data collection, while their ongoing language development makes them particularly sensitive to motivational and autonomy-related interventions (Dörnyei, 2005). This level is theoretically suitable for examining shifts in motivation and autonomy, as learners are transitioning from guided to more self-directed learning, a critical phase for developing the ideal and ought-to L2 selves (Dörnyei, & Ryan, 2015). Informing participants about the aim of the research and the nature of the groups was handled with care to minimize potential biases. Participants were informed that they were part of a study exploring different teaching methods, emphasizing the importance of maintaining unbiased responses. The specifics of being in either a control or experimental group were intentionally withheld to prevent any awareness that might influence their answers.

Instruments

Two questionnaires served as both pre-tests and post-tests, chosen for their established validity and reliability within the Iranian context, as reported in prior research.

Motivational Self System Questionnaire

This questionnaire, translated and validated by Papi (2010), consists of 12 items encompassing two subscales: six items measure learners' ideal L2 self, and six items measure learners' ought-to L2 self. Participants rated their responses on a 6-point Likert scale ranging from «strongly disagree» to «strongly agree». Reliability and validity of this questionnaire have been analyzed and reported in previous studies in Iran (e.g., Ghanizadeh, Eishabadi, & Rostami, 2015; Ghanizadeh & Rostami, 2015; Papi, 2010; Rostami, Ghanizadeh, & Ghonsooly, 2015). In the present study, Cronbach's reliability for ideal L2 self and ought-to L2 self were .82 and .91, respectively, indicating high reliability.

Learner Autonomy Questionnaire

The Learner Autonomy Questionnaire, designed and validated by Zhang and Li (2004), was used as both pre-test and post-test to assess participants' autonomy in learning English as a foreign language. Comprising 21 items rated on a five-point scale from 'Never' to 'Always', this questionnaire has demonstrated high content validity and reliability in the Iranian context (Nematipour, 2012). In this study, Cronbach's alpha reliability coefficient for this questionnaire was .75, indicating acceptable reliability. The questionnaire measures multiple dimensions of learner autonomy, including cognitive strategies (e.g., setting learning goals and selecting appropriate resources), behavioral engagement (e.g., initiating and sustaining learning tasks independent-

ly), and metacognitive planning (e.g., self-monitoring and evaluating progress).

Procedure

The researchers ensured the ethical considerations of the study, including informed consent and voluntary participation. The research was conducted over several weeks in a private language institute between September and October 2019, with two 1.5-hour sessions per week. It is worth mentioning that, to address ethical concerns, the study adhered to rigorous ethical standards by ensuring equal instructional content and time for both groups. Importantly, the researchers provided this course to students for free, covering the financial costs associated with study, emphasizing their dedication to a fair and unbiased investigation. The following steps were undertaken to conduct this research.

Homogenization of the Participants

The participant selection process began with the administration of the Quick Placement Test (QPT) as a placement test, targeting upper-intermediate students. Developed by Oxford University Press and Cambridge English for speakers of other languages (ESOL), the QPT is a valid and efficient English language proficiency test widely used for placement purposes. Forty-seven female students participated initially, but six were excluded as their score did not fall within the upper-intermediate range, leaving 41 homogenized participants. According to designers of the QPT test, participants who scored between 40 and 47 on the test are upper-intermediate. These homogenized participants were randomly assigned into either the experimental or control group with the same teacher assigned to both groups to minimize the teacher's potential impact on the results.

Control Group

In the control group, motivation and autonomy levels were assessed before the treatment via pre-test questionnaires. Both groups received identical instruction in reading and writing skills, with teacher-made tests administered for each session. The key distinction between the control and experimental groups was the assessment process. In the control group, students received grades for their test performance, accompanied by teacher-made tests with ten different topics, each requiring approximately 100 words of written response. However, these tests were evaluated using grades alone, without descriptive comments. Post-test questionnaires were administered after ten sessions.

Experimental Group

In the experimental group, pre-test questionnaires were used to assess motivation and autonomy levels. For this group, all 21 students received written descriptive comments, along with an opportunity for self-correction, on

teacher-made tests for each session. FIT was applied through the SE2R model, involving “summarize, explain, re-direct and resubmit” (Barnes, 2015). This approach aimed to provide comprehensive feedback in line with the mastery learning approach. Implementing these four stages, the teacher first began by summarizing student work. This assist situates both learners and instructor when they refer to that piece of feedback; it describes the assignment that was attempted. Next comes the explanation. Here, the teacher provided evidence to explain the level of mastery that the students had achieved (Barnes, 2015). The level of mastery that students achieved depends on their demonstration of the specified learning outcomes. The instructor in this phase developed these results in collaboration with the students by allowing them time to reflect and discuss these outcomes with their peers and then re-wording them in more student-friendly language. When both the student and instructor agreed about the learning results, the instructor provided evidence directly from some parts of the work to mention which outcomes were met, which were not, and which need improvement. When the learners mastered all the learning objectives, then they did not proceed to the re-direction step. When the instructor believed that their work needed improvement, they offered guidance or recommendations or prompts which encouraged the learners to work on special parts of tasks which do not yet show mastery. The instructor provided this redirection feedback to the learners, and they had a chance to revise their work with particular attention to the highlighted details. In the last phase, students resubmitted their work. To be clear, the final descriptive feedback using Mark Barnes’ SE2R model included a sentence at the end which asked the learner to resubmit their work, once it has been revised (Clark, 2017). As stated before, in the experimental group, students’ examinations were assessed using descriptive feedback by Mark Barnes’ SE2R model which included a sentence instead of presenting grades. For example, the teacher wrote: “Please resend me when you’ve made the changes I asked you, so I can return to your text and re-evaluate it for you.” Finally, after ten sessions, all 21 students of this group received the same motivation and autonomy questionnaires as post-tests to compare with the pre-test.

Data Analysis

Statistical Package for Social Sciences (SPSS 24) was used for analyzing data and computing normality, descriptive statistics, reliability analysis of the instruments, and independent-sample t-test. An independent sample t-test was used to make sure that the learners were at the same level of motivation and autonomy before the treatment. The homogeneity of variances for each combination of the groups was checked using Levene’s test. To address the research questions concerning motivation and autonomy scores after treatment, another independent-sample t-test was employed.

Design

This study is categorized as quasi-experimental due to non-random selection of participants based on the QPT and random assignment into experimental and control groups. Both groups participated in pretest and post-test assessments. Dependent variables encompass students’ ideal and ought-to-I2-selves’ motivation and autonomy, while the independent variable involves teachers’ adoption of the gradeless approach through descriptive feedback.

RESULTS

The Shapiro-Wilk test confirmed normal data distribution, with p-values above .05 for both groups in motivation and autonomy questionnaires. This indicates the appropriateness of employing parametric tests. The assumptions of the t-test were checked for both pre-test and post-test scores. Table 1 presents the descriptive statistics for both groups in the pre-test assessments of motivation and autonomy.

Table 1 shows the baseline pre-test data for 41 Iranian EFL learners across the experimental (n=21) and control (n=20) groups, which presents mean scores and standard deviations for the variables of ideal L2 self, ought-to L2 self, and autonomy. Specifically, the experimental group recorded a pre-test mean of 27.66 (SD = 3.91) for ideal L2 self and 27.38 (SD = 3.39) for ought-to L2 self, while the control group reported means of 26.90 (SD = 4.47) and 25.30 (SD = 4.29) respectively; for autonomy, the experimental group’s pre-test mean was 52.76 (SD = 9.75) compared to 56.40 (SD = 9.65) for the control group.

Pre-Test Comparisons: Baseline Equivalence

To determine the significance of these differences, independent sample t-tests were conducted with results presented in Table 2.

As outlined in Table 2, no significant differences were observed between the groups concerning ideal L2 self ($t = -.58$, $p = .56$), ought-to L2 self ($t = -1.72$, $p = .09$) and autonomy ($t = 1.20$, $p = .23$) during the pre-test phase. This suggests that, with a 95% confidence interval, there were no significant differences between the mean scores of the control and experimental groups in terms of ideal L2 self, ought-to L2 self, and autonomy prior to the treatment.

Post-Test Differences: Treatment Effects

In order to answer the research questions, another independent sample t-test was used. Table 3 presents the descriptive statistics for both groups in the post-test assessments of motivation and autonomy.

The table presents the post-test data for the experimental and control groups, consisting of 41 Iranian EFL learners

Table 1
Descriptive Statistics of Two Groups in Motivation and Autonomy in the Pre-Test Assessments

Variable	Group	N	Mean	Std. Deviation	Std. Error Mean
Ideal L2 Self	Control	20	26.90	4.47	.99
	Experimental	21	27.66	3.91	.85
Ought-to L2 Self	Control	20	25.30	4.29	.95
	Experimental	21	27.38	3.39	.74
Autonomy	Control	20	56.40	9.65	2.15
	Experimental	21	52.76	9.75	2.12

Table 2
Results of the Independent Sample t-test of Motivation and Autonomy in the Pre-Test Assessments

Variable	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Ideal L2 Self	-.58	39	.56	-.76	1.31
Ought-to L2 Self	-1.72	39	.09	-2.08	1.20
Autonomy	1.20	39	.23	3.63	3.03

Table 3
Descriptive Statistics of Two Groups in Motivation and Autonomy in the Post-Test Assessments

Variable	Group	N	Mean	Std. Deviation	Std. Error Mean
Ideal L2 Self	Control	20	27.15	4.54	1.01
	Experimental	21	30.38	2.97	.64
Ought-to L2 Self	Control	20	28.90	4.02	.90
	Experimental	21	25.66	2.65	.57
Autonomy	Control	20	55.10	8.96	2.00
	Experimental	21	60.76	8.61	1.88

(experimental: n=21; control: n=20), across the variables of ideal L2 self, ought-to L2 self, and autonomy. For ideal L2 self, the experimental group recorded a post-test mean of 30.38 with a standard deviation of 2.97, while the control group reported a post-test mean of 27.15 with a standard deviation of 4.54. For ought-to L2 self, the experimental group’s post-test mean was 25.66 with a standard deviation of 2.65, compared to 28.90 with a standard deviation of 4.02 for the control group. Regarding autonomy, the experimental group exhibited a post-test mean of 60.76 with a standard deviation of 8.61, whereas the control group’s post-test mean was 55.10 with a standard deviation of 8.96.

Effect Size and Interpretation

To determine the significance of these differences, independent sample t-tests were conducted with results presented in Table 4. Eta squared was used to check the effect size.

As presented in Table 4, post-test comparisons revealed significant differences between the experimental (grade-less) and control (graded) groups across all measured outcomes. For ideal L2 self-motivation, the experimental group outperformed the control group ($t = -2.70, p = .01, \eta^2 = .15$), indicating a large effect size (Cohen, 1988). This

Table 4

Results of the Independent Sample T-Test of Motivation and Autonomy in the Post-Test Assessments

		t	df	Sig (2-tailed)	Mean Difference	Std. Error Difference
Post-test	Ideal L2 Self	-2.70	39	.01	-3.23	1.19
	Ought-to L2 Self	3.05	39	.00	3.23	1.05
	Autonomy	-2.06	39	.04	-5.66	2.74

suggests a substantial pedagogical impact, as the gradeless approach markedly enhanced students' intrinsic aspirations to achieve language proficiency. For ought-to L2 self-motivation, the control group showed higher scores ($t = 3.05$, $p = .00$, $\eta^2 = .19$), also reflecting a large effect size, implying that traditional grading reinforced extrinsic obligations more effectively, though potentially at the cost of intrinsic motivation. For autonomy, the experimental group demonstrated greater improvement ($t = -2.06$, $p = .04$, $\eta^2 = .09$), with a medium effect size, indicating a meaningful increase in students' ability to self-direct their learning processes. These findings underscore the differential impact of gradeless assessment, involving descriptive feedback and self-correction, in fostering intrinsic motivation and autonomy while reducing reliance on external pressures in the Iranian EFL context

a collectivist Iranian context. In addition, the statistically significant increase in autonomy aligns with Lamb's (2010) concept of assessment for autonomy. He defined assessment for autonomy as any assessment for which the priority in its design and practice is to serve the purpose of promoting pupils' autonomy. Jácome (2012) also reported similar findings emphasizing the role of teacher-student partnership assessment using feedback and self-evaluating skills in promoting learner autonomy. Furthermore, according to Lamb (2010), providing feedback not only enhances students' language learning but also their autonomy and motivation. He also continued that to enable students to become more autonomous learners, teachers should encourage them to involve not only what they have learned, but also how they learned it through a meaningful assessment such as providing feedback.

DISCUSSION

Main Findings and Theoretical Interpretations

In this study, the researchers conducted experimental research to examine the impact of implementing the gradeless method on Iranian EFL students' ideal and ought-to L2-selves motivation and autonomy. Grounded in theoretical considerations, the hypothesis suggested that providing descriptive feedback without grades would maintain or enhance students' motivation and autonomy. The findings revealed a significant increase in ideal L2 self-motivation and autonomy, alongside an unexpected decrease in ought-to L2 self-motivation, offering insights into the psychological effects of assessment practices in the Iranian EFL context. This study also found that providing numerical grades enhances ought-to L2 self-motivation at the expense of ideal L2 self-motivation. These results directly address the research questions by highlighting how gradeless assessment reshapes motivational constructs. Drawing on the L2 Motivational Self System (L2MSS), the rise in ideal L2 self reflects learners' strengthened internalized aspirations, as theorized by Dörnyei (2005), particularly when external grading is removed. Self-Determination Theory (SDT) further supports this, indicating that descriptive feedback enhances competence and relatedness, boosting intrinsic motivation. This extends Wu et al. (2014), who linked task-involving environments to motivation, by demonstrating these effects in

Unexpected Findings and Explanations

Regarding the impact of the type of assessment on students' motivation, the findings align with FIT and support previous research across various contexts (e.g., Butler & Nisan, 1986; Lipnevich & Smith, 2008; Stefanou & Parkes, 2003; Taylor & Nolen, 2008; Wu et al., 2014). The findings suggest that when students receive constructive feedback focused on improvement rather than grades, they are more likely to be driven by a genuine desire to attain their ideal language proficiency. Wu et al. (2014) emphasized the close relationship between classroom assessment and motivation, highlighting the positive impact of task-involving environments, diverse assessment tasks, descriptive feedback, and clear learning objectives on intrinsic motivation. However, providing a competitive classroom environment, instructing based on the test, and providing students grades are associated with extrinsic motivation. In line with their findings, our study supports the idea that creating a task-involving classroom environment, and providing descriptive feedback enhances students' ideal L2-self motivation. Therefore, the results of our study contribute to the existing body of knowledge by providing additional evidence of the positive impact of gradeless assessment using descriptive feedback on students' motivation, aligning with the principles highlighted by Wu et al. (2014). Additionally, Stefanou and Parkes (2003) highlighted the potential impact of the grading system on learner behavior and motivation, and

our findings substantiate this claim. Specifically, our results emphasize the significant effect of assessment types on motivation, reinforcing the importance of adopting assessment practices that prioritize descriptive feedback over grades. In connecting our findings to the broader literature, we corroborate the view that an overemphasis on grades, as indicated in prior studies (Coffield, 2012; Matthews & Noyes, 2014; McMorran & Ragupathi, 2020), indeed leads to detrimental impacts on the learning process. The present study further validated that learners who engage in self-correction assessments exhibit higher levels of motivation. Stiggins (2005) mentioned that “If students play even a small role in setting the (learning achievement) target...we can gain considerable motivational and therefore achievement benefits” (p.244). Therefore, teachers can increase students’ motivation by engaging them in the process of assessment, which can be done by the self-correction assessment method. Stiggins (2005) also emphasized the importance of self-correction which increases students’ responsibility for improving their sense of control over their success and accordingly enhances their motivation. The decrease in the mean score of ought-to L2 self-motivation in the experimental group from 27.38 in the pre-test to 25.66 in the post-test warrants deeper analysis. This unexpected finding might be attributed to several factors that could have influenced participants’ perceptions and expectations during the study. Potential reasons could include variations in how students interpreted or internalized descriptive feedback, the nature of the self-correction process, or even individual differences in response to the gradeless assessment method. The concept of ought-to L2 self, as derived from Dörnyei’s L2 Motivational Self System model, represents the traits learners believe they should possess due to various duties, obligations, or expectations. While the decrease in ought-to L2 self-motivation may initially seem counterintuitive, it aligns with the theoretical perspectives suggesting that traditional grading systems, emphasizing extrinsic rewards and external obligations, may negatively impact intrinsic motivation (Lee, Lallie, & Michaelides, 2023). This finding demonstrates a potential trade-off in gradeless systems, where increased autonomy and intrinsic drive (ideal L2 self) may reduce compliance-based motivation (ought-to L2 self). Using L2MSS, the de-emphasis on grades likely weakens the ought-to self, which thrives on external pressures. SDT complements this, suggesting that removing extrinsic regulation may destabilize ought-to motivation unless supported by autonomy-enhancing practices. In Iran’s collectivist culture, this decline may reflect reduced societal or familial expectations, though our data lacks direct evidence to confirm this, highlighting a need for further investigation.

Implications for EFL Assessment Practices

This section discusses these findings in light of existing literature, addresses their cultural implications, and considers potential limitations and adverse effects of the gradeless approach. Our findings imply that gradeless assessment,

with its focus on descriptive feedback, can enhance EFL learning by boosting ideal L2 self and autonomy, challenging traditional grading paradigms. This extends Butler and Nisan (1986), who highlighted feedback’s role in motivation, by showing its efficacy in a non-Western context. Practically, educators might integrate self-correction to foster agency, though balancing autonomy with accountability remains critical. This study suggests self-correction as a valid proxy for learner agency, offering a novel approach to assessment design. At the institutional level, language programs could explore integrating gradeless methods to create environments that support sustainable motivation, though careful implementation is needed to address diverse learner needs.

Cultural Considerations

The findings put forward several implications for curriculum design, teacher training, and assessment policy, particularly in contexts seeking to foster sustainable motivation and independent learning. Language instructors are encouraged to adopt descriptive feedback practices to nurture students’ intrinsic motivation and self-directed learning. Teacher training programs should emphasize formative assessment techniques, equipping educators to implement feedback-driven approaches effectively. In the Iranian EFL context, where traditional grading often limits learner agency, the gradeless method’s focus on self-correction likely fostered metacognitive strategies, aligning with Oxford’s (2011) framework for autonomous learning. However, constructs like autonomy and the ought-to L2 self are culturally embedded, and their interpretation may differ in other educational systems. For instance, in individualistic cultures like Scandinavian countries, where autonomy is already emphasized, the impact of gradeless assessment may be less pronounced (Hofstede, 2011). Conversely, in assessment-driven East Asian systems, students may rely more heavily on grades for motivation, potentially limiting the efficacy of descriptive feedback (Cheng & Curtis, 2010). The Iranian context, characterized by collectivist values, likely shapes how gradeless assessment impacts motivation, with the decline in ought-to L2 self reflecting a reduced emphasis on external approval, a cultural norm in collectivist settings.

Limitations and Directions for Future Research

While the gradeless approach yielded positive outcomes for ideal L2 self-motivation and autonomy, potential negative effects must be considered. The decrease in ought-to L2 self-motivation suggests that some students, particularly those reliant on external validation, may experience uncertainty or reduced motivation without grades. High-performing students accustomed to competitive environments may find the lack of grades demotivating, as grades often serve as a benchmark for achievement (Pulfrey et al., 2011). Additionally, less confident learners may feel anxious without the clear structure of numerical feedback, as noted by Sackstein (2015), who observed challenges for high-achieving

ing students transitioning to gradeless systems. In the Iranian context, where grades are deeply tied to academic and social status, this shift could exacerbate uncertainty for some learners. Research indicates that gradeless learning can lead to uncertainty and identity loss among high-performing students accustomed to competitive environments, potentially destabilizing extrinsic motivation (Kjærgaard, et al., 2023). Furthermore, learners in collectivist cultures may struggle to adapt to feedback-driven systems without clear performance metrics, which can exacerbate anxiety and reduce motivation (McNall & Gravelin, 2024). These findings suggest that while gradeless systems have potential, their success depends on careful implementation and cultural adaptation. Several limitations affect this study's generalizability. The cultural specificity to Iranian EFL learners, a sample size of 41, and the short-term intervention without follow-up data constrain the findings. Reliance on self-report instruments may also introduce bias. Future research should employ longitudinal designs to assess long-term effects, use mixed-methods for deeper insights, and compare outcomes across diverse educational cultures to address these gaps and refine gradeless assessment practices.

CONCLUSION

This study explored the impact of gradeless assessment, utilizing descriptive feedback and self-correction, on the motivation and autonomy of Iranian EFL learners. The findings suggest that descriptive feedback and self-correction practices can significantly enhance learners' ideal L2 selves and foster greater autonomy. These preliminary results underscore the potential of formative assessment to reorient lan-

guage education toward meaningful learning rather than grade-driven outcomes. Looking forward, these findings pave the way for further exploration of alternative assessment models in language education. Future research should investigate the long-term effects of gradeless assessment across diverse cultural and proficiency contexts to enhance its applicability. By building on these insights, educators and researchers can advance learner-centered pedagogies that foster motivation and autonomy, contributing to the evolution of EFL assessment practices globally. This study offers new insights into gradeless assessment by demonstrating its potential to enhance L2MSS dimensions in under-researched collectivist contexts like Iran, while revealing the challenge of balancing autonomy with accountability in assessment systems.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Soroor Rostami: conceptualization; formal analysis; investigation; methodology; resources; writing – original draft; writing - review and editing.

Mansooreh Hosseinnia: conceptualization; investigation; methodology; resources; writing – original draft; writing - review and editing.

REFERENCES

- Barnes, M. (2015). *Assessment 3.0: Throw out your gradebook and inspire learning*. Corwin.
- Belton, T. R. (2022). *The impact of grading systems on middle school students' summative mathematics achievement test scores* [Unpublished doctoral dissertation]. Regent University.
- Brookhart, S. M. (2008). *How to give effective feedback to your students*. Association for Supervision and Curriculum Development.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58(1), 1-14. <https://doi.org/10.1111/j.2044-8279.1988.tb00874.x>
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210-216. <https://doi.org/10.1037/0022-0663.78.3.210>
- Cain, J., Medina, M., Romanelli, F., & Persky, A. (2022). Deficiencies of traditional grading systems and recommendations for the future. *American Journal of Pharmaceutical Education*, 86(7). <https://doi.org/10.5688/ajpe8850>
- Cheng, L., & Curtis, A. (2010). *English language assessment and the Chinese learner*. Routledge.
- Clark, M., R. (2017). *Reducing the use of and emphasis on grades in high school physics classes* [Unpublished MA thesis]. University of Toronto.
- Coffield, F. (2012). To grade or not to grade. *Adult Learning*, 23(4), 38-39.
- Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge University Press.

- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.
- Ghanizadeh, A., Eishabadi, N., & Rostami, S. (2015). Motivational dimension of willingness to communicate in L2: The impacts of criterion measure, ideal L2 self, family influence, and attitudes to L2 culture. *International Journal of Research Studies in Education*, 4(5), 13-24. <https://doi.org/10.5861/ijrse.2015.1261>
- Ghanizadeh, A., & Rostami, S. (2015). A Dörnyei-inspired study on second language motivation: A Cross-comparison analysis in public and private contexts. *Psychological Studies*, 60(3), 292-301. <https://doi.org/10.1007/s12646-015-0328-4>
- Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: A systematic review and meta-analysis. *Educational Research Review*, 34, 100401. <https://doi.org/10.1016/j.edurev.2021.100401>
- Guskey, T. R., & Lee, A. J. (2013). *Answers to essential questions about standards, assessments, grading, and reporting*. Corwin.
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture*, 2(1), 8. <https://doi.org/10.9707/2307-0919.1014>
- Jácome, E. P. (2012). Promoting learner autonomy through teacher-student partnership assessment in an American high school: A cycle of action research. *Profile Issues in Teachers Professional Development*, 14(2), 145-162.
- Jaschik, S. (2009). *Imagining college without grades*. Inside Higher Education.
- Kjærgaard, A., Mikkelsen, E. N., & Buhl-Wiggers, J. (2023). The gradeless paradox: Emancipatory promises but ambivalent effects of gradeless learning in business and management education. *Management Learning*, 54(4), 556-575. <https://journals.sagepub.com/doi/pdf/10.1177/13505076221101146>
- Kluger, A., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 2, 254-284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Lamb, T. (2010). Assessment of autonomy or assessment for autonomy? Evaluating learner autonomy for formative purposes. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 98-119). Multilingual Matters.
- Lee, D., Lallie, H. S., & Michaelides, N. (2023). The impact of an employee's psychological contract breach on compliance with information security policies: Intrinsic and extrinsic motivation. *Cognition, Technology & Work*, 25(2), 273-289. <https://doi.org/10.1007/s10111-023-00727-5>
- Lerner, R. E., Grolnick, W. S., Caruso, A. J., & Levitt, M. R. (2022). Parental involvement and children's academics: The roles of autonomy support and parents' motivation for involvement. *Contemporary Educational Psychology*, 68, 102039. <https://doi.org/10.1016/j.cedpsych.2021.102039>
- Lim, K. (2024). Assessing beyond grades: Unravelling the implications on student learning and engagement in higher education. *Assessment & Evaluation in Higher Education*, 49(5), 665-679. <https://doi.org/10.1080/02602938.2024.2314703>
- Linder, A., Gerdtham, U. G., & Heckley, G. (2023). *Adolescent mental health: Impact of introducing earlier compulsory school grades*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4499202
- Lipnevich, A. A., & Smith, J. K. (2008). *Response to assessment feedback: The effects of grades, praise, and source of information*. Educational Testing Service.
- Marquez, E., & Garcia, S. (2023). Keeping calm and staying balanced: exploring the academic pressures faced by engineering students to attain high grades and their impact on mental health. *Paper presented at 2023 ASEE Annual Conference & Exposition, Baltimore, Maryland*. <https://peer.asee.org/43908>
- Matthews, R., & A. Noyes. (2014). To grade or not to grade: Balancing formative and summative assessment in Post-16 teacher trainee observations. *Journal of Further and Higher Education*, 40(2), 247-261. <https://doi.org/10.1080/0309877x.2014.953456>
- McMorran, C., & Ragupathi, K. (2020). The promise and pitfalls of gradeless learning: Responses to an alternative approach to grading. *Journal of Further and Higher Education*, 44(7), 925-938. <https://doi.org/10.1080/0309877x.2019.1619073>
- Mendoza, N. B., Yan, Z., & King, R. B. (2023). Supporting students' intrinsic motivation for online learning tasks: The effect of need-supportive task instructions on motivation, self-assessment, and task performance. *Computers & Education*, 193, 104663. <https://doi.org/10.1016/j.compedu.2022.104663>
- Mortazavizadeh, H., Nili, M., Nasr Isfahani, A., & Hassani, M. (2017). Teachers' lived experiences about teaching-learning process in multi-grade classes. *Journal of Education and Learning*, 6(4), 354-363. <http://doi.org/10.5539/jel.v6n4p354>
- Muho, A., & Taraj, G. (2022). Impact of formative assessment practices on student motivation for learning the English language. *International Journal of Education and Practice*, 10(1), 25-41. <https://doi.org/10.18488/61.v10i1.2842>
- McNall, L. A., & Gravelin, C. R. (2024). An exploratory study of student perceptions of ungrading in upper-level psychology courses. *Teaching of Psychology*, 00986283241293413. <https://doi.org/10.1177/00986283241293413>

- Nematipour, M. (2012). A study of Iranian EFL learners' autonomy level and its relationship with learning style. *English Linguistics Research*, 1(1), 126-136. <https://doi.org/10.5430/elr.v1n1p126>
- Oxford, R. L. (2011). Strategies for learning a second or foreign language. *Language Teaching*, 44(2), 167-180. <https://doi.org/10.1017/S0261444810000492>
- Papi, M., (2010). The L2 motivational self system, L2 anxiety, and motivated behavior: A structural equation modeling approach. *System*, 38, 467-479. <https://doi.org/10.1016/j.system.2010.06.011>
- Pishghadam, R., Hashemi, M. R., & Kermanshahi, P. N. (2011). Self-correction among Iranian EFL learners: An investigation into their preferences for corrective feedback. *Journal of Language Teaching & Research*, 2(5), 957-962. <https://doi.org/10.4304/jltr.2.5.957-962>
- Pratomo, H. W., & Kuswati, Y. (2022). The effect of teacher motivation on student achievement in Islamic senior high school. *International Journal of Educational Qualitative Quantitative Research*, 1(2), 16-22. <https://doi.org/10.58418/ijeqqr.v1i2.17>
- Pulfrey, C., Buchs, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: the mediating role of autonomous motivation. *Journal of Educational Psychology*, 103, 683-700. <https://doi.org/10.1037/e549962013-069>
- Quesada, L. A., Oviedo, F. C., Nunez, S. V., Charpentier, K. A., & Atencio, M. F. Z. (2025). Extrinsic and intrinsic motivation and its impact on students' performance in the EFL classroom. *Ciencia Latina Revista Científica Multidisciplinar*, 9(1), 5670-5682. https://doi.org/10.37811/cl_rcm.v9i1.16244
- Razavi, A. (2016). Three components of descriptive evaluation: Feedback, self-assessment, and peer-assessment. *Journal of Instruction and Evaluation*, 9(36), 49-74.
- Rojstaczer, S., & Healy, C. (2012). Where A is ordinary: the evolution of American college and university grading, 1940-2009. *Teachers College Record*, 114(7), 1-23. <https://doi.org/10.1177/016146811211400707>
- Rostami, S., Ghanizadeh, A., & Ghonsooly, B. (2015). External factors affecting second language motivation: The role of teacher burnout and family influence. *Iranian Journal of Applied Linguistics*, 18(2), 165-187. <https://doi.org/10.18869/acadpub.ijal.18.2.165>
- Sackstein, S. (2015). Hack learning: Hacking assessment. Cleveland Times 10.
- Smith, Z. R., Flax, M., Becker, S. P., & Langberg, J. (2023). Academic motivation decreases across adolescence for youth with and without attention-deficit/hyperactivity disorder: Effects of motivation on academic success. *Journal of Child Psychology and Psychiatry*, 64(9), 1303-1313. <https://doi.org/10.13056/acamh.24604>
- Stefanou, C., & Parkes, J. (2003). Effects of Classroom assessment on student motivation in fifth-grade science. *The Journal of Educational Research*, 96(3), 152-162. <https://doi.org/10.1080/00220670309598803>
- Stiggins, R. J. (2005). *Student-involved assessment for learning* (5th ed.). Merrill/Prentice Hall.
- Supiano, B. (2019). *Grades can hinder learning. What should professors use instead*. Chronicle of Higher Education.
- Tavallali, E., & Marzban, A. (2015). Becoming autonomous learners through self-regulated learning. *Journal of Applied Linguistics and Language Research*, 2(3), 72-83
- Taylor, C. S. & Nolen, S. B. (2008). *Classroom assessment: Supporting teaching and learning in real classrooms* (2nd ed.). Pearson Education, Inc.
- Wu, Y., Cheng, L., & Bettney, E. (2014). Exploring the relationships between assessment and motivation: Perspectives from teacher candidates. *Paper presented at the 58th International Council on Education for Teaching World Assembly, University of Ontario Institute of Technology, Oshawa, Ontario, Canada.*
- Zedan, R. (2021). Student feedback as a predictor of learning motivation, academic achievement, and classroom climate. *Education and Self-Development*, 16(2), 27-46. <https://doi.org/10.26907/esd.16.2.03>
- Zhang, L.X., & Li, X. X. (2004). A comparative study on learner autonomy between Chinese students and west European students. *Foreign Language World*, 4, 15-23.