

Reforms in High-Stake University Entrance Examinations in Iran: A Critical 'Base-line' Study

P. Shayestefar¹, G.R. Kiany², R.G. Samar³, R. Akbari⁴,

Received: 2011/10/4

Accepted: 2012/4/14

Abstract

A rich research base suggests that high-stakes tests reforms serve as vehicles for promoting quality of learning, standards of teaching, and credible forms of accountability. Iran's last decade policies to reform its University Entrance Examinations (UEEs) heralded such a case. What sparked off this reform was the long-lasting pernicious impact of the UEEs on curriculum, instruction, learning, and on societal values and access to the upward social mobility. However, attempts to introduce intended changes are often not as effective as their planners hoped. A scrutiny of change 'antecedent conditions' as well as its initial 'process of diffusion' i.e., a 'baseline study' will ameliorate such a failure (Weir & Roberts, 1994, Fekete *et al*, 1999). Assessing the feasibility of the UEEs reform through such a 'baseline study' is a gap in Iran's reform initiatives. As such, the authors applied Henrichsen's (1989) Hybrid Diffusion Model (HDM), underscoring an awareness of and a need for evaluation of any changed program from a multiplex of factors, as its theoretical framework to critically evaluate the reformed program. The paper thus first presents the policy deliberations and steering National Documents that pushed through an assessment-led reform in Iran's education. Then it sketches a detailed discussion of the contextual policies and practices of the long-lasting UEEs program, its alternative and associated stakes. While the results highlight key political dynamics which drove national policymaking, they are suggestive of the challenges, controversies, and risks that thwart the success rate of *ideals* intended by the underpinning policies.

Keywords: High-Stakes Tests; Assessment-Led Reform; Intended Consequences; Students' Background Records.

-
1. Ph.D. Student, Department of English, Tarbiat Modares University, parishayeste@yahoo.com
 2. Associate Professor of Applied Linguistics, Department of English, Tarbiat Modares University, Rezakiany@yahoo.com
 3. Assistant Professor of Applied Linguistics, Department of English, Tarbiat Modares University, rgsamar@modares.ac.ir
 4. Assistant Professor of Applied Linguistics, Department of English, Tarbiat Modares University, akbari_r@modares.ac.ir

1. Introduction

Over the past four decades, Iran has taken several piecemeal efforts to modify its educational system and improve the quality of its outcomes. Although some of these endeavors have been integrative and occasionally overlapping, their effects did not always lead into substantial improvement but may simply into reordering or superficial modifications of the *status quo*. As such, a litany of criticisms provoked by relative dissatisfactions with schooling quality, outcomes, and outcome-measurement system on the one hand, and a leap for lining up with globalization as ‘a transition into third millennium’ (Water, 1995) on another hand, crystallized an essential reform mission for the country to embark on. The general *reform wave* that has swept over the country by the policy deliberation councils since 2001’s academic debates patently portray reform initiatives in many areas of education from policy to practice, from curriculum to assessment.

To address public concerns about the quality of education and its outcomes, the turn of the century in Iran, thus, heralded radical policy-making enterprises to opt for *learner’s active, life-long, constructive*

meaningful learning, also for an *integration of constructivism* in both *teaching pedagogy* and *assessment of students’ learning in its new conception*. (Iran’s 20-Year Vision, 2005; The Document of National Curriculum-DNC, 2010). Notably, such macro-concerns that have been largely overlooked both at general education and higher education alike for years, recently received widespread supports. These policies, in a parallel fashion, not only have drastically pushed through *curriculum reform* but also indispensably put *assessment reform* in place too. The legitimacy of simultaneity of curriculum-assessment reforms in this sense can be clearly endorsed by a rather test-driven accountability of the country that, in turn, induces diverse consequences on curriculum, instruction, and learning. Indeed, despite the allegedly balanced reciprocity between the national curriculum and the educational measurement system, the balance has tilted toward measurement or testing system due to much reliance on large-scale testing such as University Entrance Examinations (UEEs) or High school National Achievement Exams (HNAEs) as

the *stake* is so much high throughout the country particularly for the UEEs.

Given such a sensitive stake of the UEEs that cannot be seen without serious consequences on curriculum, instruction, and learning, it would not seem too much of an exaggeration to say that tests-based reform as a far-reaching and high-stake policy is a prime candidate for promoting the quality of curriculum, teaching, and learning in this country. Regarding this strong impact of the UEEs program, it is worth mentioning here that tests alteration of any type should come aligned with the intended curriculum of the country. Suffice here to mention that necessary taskforces for reinventing and restructuring of the existing curriculum were formed in 2006, and finally, painstaking endeavors resulted into the development of the DNC which has been recently finalized after its 10th revision. As an important steering documents of education besides the three others (i.e., The Philosophy of Education in the Islamic Republic of Iran; Fundamental Principles in Islamic Education; and The Roadmap of the Official and General Educational System-ROGES, 2009), the DNC has enjoyed serious alterations in

content, focus, organization, and time allocation for its content, and components skills. The reformed curriculum trounces the flaws of the existing practiced and learned curriculum; integrates challenging academic contents, skills, and achievement standards; congregates the technology era's emerging needs, national identity, local realities, and interests; and corrects for any disparity known to foster rote-, or mechanical learning.

Main drives for these curricular and structural alterations, as its proponents called, are traced in theoretical and value-laden foundations of religious education; local and national potentials; successful and thriving experiences of other nations' education; innovations in curricular approaches and methods; intention to plan a comprehensive science map for the country as part of the "*cultural engineering*" in general education; and finally needs and new demands of the country (DNC, 2010: 6-7). More prominently, along with such vein of reform-led movements, the recent 6-year debates on tests-driven accountability ushered in a fundamental reform the UEEs program which has over decades been boosted by a chronic testing hegemony.

Such a coincidence of curriculum and assessment reforms in the past six years will have the potential to impede 'curriculum shrinkage' observed in our education for more than four decades. In clear words, as long as the UEEs are used for important decision makings about university entrance, they have immense importance for the people and institutions involved so that they will continuously endure narrowing the reformed curriculum as before. In a recent review of literature, a number of researchers (see for e.g., Decker & Bolt, 2008; Kikuchi & Browne, 2009; Munoz & Alvarez, 2010; Wall & Horak, 2008) show how high-stake tests of this kind will bring about a more equal curriculum by making a renewed focus on what is measured by that test. Part of such debates over 'test washback' or 'test impact' derives from negative consequences of high-stakes tests on the curriculum, teachers, learners, and other stakeholders involved. Nevertheless, a positive picture of consequences has been also reported in this literature.

Inspired by such an argument threads, proponents of testing-based reforms in the country in a lively collaboration with those critics of the UEEs (e.g, Kia &

Bozorgi, 2006; Shojaee & Gholipour, 2005; Hajforoush, 2002) persuasively recommended for a policy remedy to supplant the One-shot centralized gate-keeper UEEs with a more sound and standard measure such as NAEP- National Assessment Educational Progress tests practiced in many countries (letters to the Research Center of the Parliament-RCP, 2004-2009) that mirrors consecutive practices and performances within high school contexts over years of schooling. Out of this educational venture, emerged a national proposal for student selection whose official realization and application not only brings an alignment focus on the standards of the new curriculum but as has been rightly put forward by some scholars (e.g., Black and Williams, 1998; Darling-Hammond, Aness, & Falk, 2001; Kornhaber, 2004; Wiggins, 1998) no longer confines students' performances only to the context of a single test like the apparently flawed UEEs.

Hence, in an attempt to wipe out and guard against a wide array of negative consequences that the UEEs have provoked, to enhance students' learning, or to remove the shackles that promote the passive rote-learning spirit as an inherent consequence

of the UEEs' structures, contents, and items, 'The Act of Student Admission to Universities' was passed by the Parliament's policymakers in 2007. According to this Act, students' academic/schooling record will replace the UEEs scores. This record is an aggregation of their high school GPAs-cumulated from their successive performances on four HNAEs (High school National Achievement Examinations) together with any extra-curricular activities, national or international records, awards, etc. The Act will come into its entire force as part of the statewide accountability in 2012-13; nonetheless, its primary implementation started in the mid of 2011 when only 25% of the total university entrance scores were accounted for by students' academic backgrounds in terms of their GPAs.

Needless to mention that such reform waves have not occurred without considerable controversies and challenges. Typically, even though high-stake tests are reported to function as "levers for change" (Alderson, 1986b); there is a substantial debate about whether their change will best meet the intended policies of the positive impacts or consequences. As to this, with new educational policy of the

country, an increasing social concern is raised calling seriously for an evidence-based evaluation of the reformed admission program. More rational for evaluation of such a new program comes from Brienbaum's (2007) quote that 'the road from theory to practice is a rocky one' showing the unorthodoxy in practices of the formulated policies, also from a pressing need to apply a broader evaluative framework such as Henrichsen's Hybrid Diffusion Model-HDM (1989) that underscores an awareness and evaluation of many factors (e.g. people, channel, and process involved) interplaying in a given changed context. Working on innovation and change theories, HDM is copious enough to assess the feasibility and the amount of risks of a new test or assessment mode through evaluating *the antecedent, process, and potential consequences* features of that new test or assessment (See Graph 1).

Therefore, to meet the complex relationship between the ideals and reals, i.e., the 'ideal policies' decided upon through reforming the UEEs program and the 'real current practice's, and to scrutinize whether the ideals address the

challenges of today's educational reality, it is quite crucial that the present assessment reform be monitored not only on its outcomes/consequences but on its 'planning and initial implementation phases' (Kennedy, 1988) as well. This *a-priori* evaluation of the two phases before evaluation of the final 'completion phase' (Kennedy, 1988) of the new program not only illuminates whether this alternative facilitates or impedes the quality promotion in education and learning which is intended by policymakers, but provides evidence for the assertion that unevaluated change programs before their final completion phase may turn the programs become 'seriously threatened, or even worse thwarted' (Stoynoff, 1989: 17).

In accord with these perspectives and aims, the article first provides a telescoped account of the rise of the UEEs reform policies of Iran, exploring the incentives that stirred such an urgent alteration in the testing agenda serving the UEEs purposes. Then, taking the Henrichsen's HDM multi-dimensional framework as a vantage point, main challenges and controversies of the new admission program will be critically sketched with a serious eye on an imperative need for examining the

present 'reform-based assessment' movement as early as possible.

2. Assessment-Led Reform: Global Perspectives & Local Priorities

Parallel to curricular reform, an intense debate on reforming the general educational measurement system of the country was also sparked off in mid last decade. Given the premise that assessment is important for obtaining a picture of curricular goals and quality (Cheng, 1999; Honig, 1987; Noble & Smith, 1994b; Popham, 1987), an immediate need for altering its conventional forms and measures is significantly apparent from the commonalities found in "Outcomes Assessment, and Evaluation" sections sketched by Iran's recent national documents. More specifically, ROGES (2009: 26, 36, 42) and DNC (2010, p. 126-139) embed separate sections on assessment policies with an implied proposal for changing the current testing system of the two ministries of Education and Higher Education. Although there were some thought-provoking reports about the relative flawed measurement systems employed over the years by the two ministries, the idea of changing the

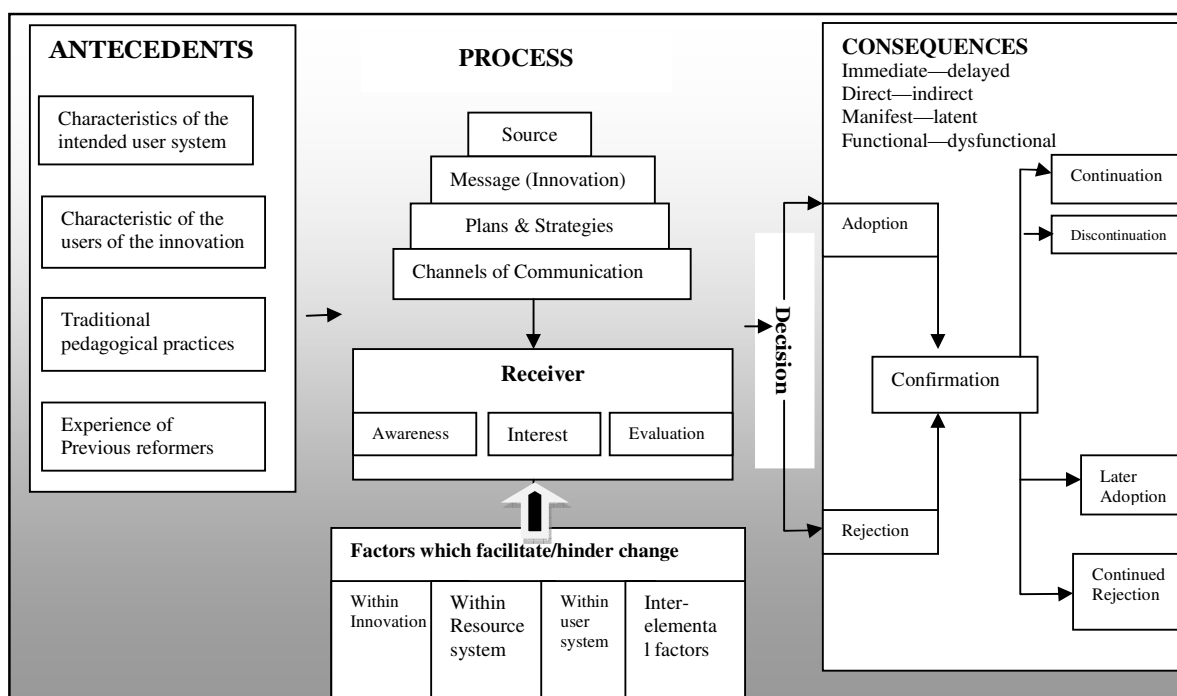


Fig 1 Henrichsen (1989) Hybrid Model of the Diffusion/Implementation Process

existing testing policies and practices for Entry criteria of the universities was not given a real formal shot before 2002 when a 3-day seminar on “*Evaluating the issues of University Entrance Examinations: criteria, measures, and concerns*” was run by a community of measurement scholars and educational experts in the State University of Esfahan (September, 2002).

What emerged out of the conference analyses on negative consequences of the UEEs, comparative studies on university Entry criteria of successful countries, and proliferation of public concerns about Iran’s

UEEs was a “Proposal for Changing the Public University Entrance Examinations” to be offered initially to the Education and Research Committee (ERP) of the Parliament of Iran (letters to the RCP, 2003). Having gone through these reports and reviews as well as having chaired several parliamentary, ministerial, and Research Councils’ meetings over two years, the Parliament passed the bill of supplanting the centralized high-stake UEEs with students’ academic background. The policy outcome of the Parliament was the “Act of Student Admission to

Universities” (Parliament, 2007) as pointed out before. Taking into account that such assessment policies require an appreciation of the socio-educational context in which the policies are introduced, modified, or abandoned (Ross, 2008), a brief description of the contextual policies and practices of the long-lasting UEEs, the alternative proposed, and the associated stakes is worth providing here.

2.1. The UEEs Hurricane in Iran

Iran has a population of 70 millions, with almost 15 million students studying at the formal Primary (5 years), Guidance/Secondary (3 years), High School (3 years), and Pre-University (1 year) education levels (in total, 50.3% girls vs. 49.7% boys). At Pre-University level, students may choose to go to higher education or less probably choose vocational/technical positions. Entry into tertiary levels, in particular into public universities, has been competitive since 1960's and got extremely competitive after Iran's revolution of 1979 when the number of candidates began to exceed universities' capacities. Yearly, over 1 million students compete for the UEE that is a Multiple-Choice test administered by

the “Sanjesh Organization” (SO)-a central testing organization for preparing, organizing, and scoring the UEEs also known for administering Standard Tests of English such as TOEFL, IELTS and GRE. Fitted to the high school fields of study (Mathematics, Natural Science, Arts and Humanities), each designed UEE consists of two sections that are administered on a three-day schedule for the three high school fields. The General-Competence section (paper) together with the Domain-Specific section (paper) of the UEE both include items that require academic knowledge of the school curricula, and are machine-scored in a month after the administrations. Candidates are essentially evaluated on the basis on their performances on these UEEs with no school academic achievement records obliged for their final admission and placement.

As the access to upward social mobility in Iran is associated with access to higher education, professionalization, and a rather-guaranteed employment, the UEEs as gate-keeping means perniciously affect the societal values and sense of competition, spreading widely through the whole society. The governmental premise

based on which the UEEs are built comes from the policies that individuals can gain equal access to higher education despite their caste, economic status, or any personal connection. However, there has been a growing critical argument against such a claim of equality provision for all members of the society (Shojaee & Gholipour, 2005). Studies have indicated the indirect effects of the family backgrounds on the selective universities admission through family's capacity to vest for the test preparation. As a rare study, Ross (2008) has sketched the current testing evolution, innovation, and policy challenges in six Asian countries (Japan, Korea, China, Thailand, Hong Kong and India) and in his overview has described how investment in after-school test coaching is increased in these countries. As a contribution to this special issue, families' investment in the UEEs preparation in Iran has also become a massive enterprise that could exemplify challenges to test fairness and equal access to higher education and mobility.

Apart from such equality debates, a great deal of discourse exists, both professional and political, regarding dissatisfaction with the UEEs program.

The seriousness of consequences of such sensitive tests raises the sense that it is the test uses and test scores interpretations that, to a large extent, determine the real or perceived positive or negative effects on various stakeholders (e.g., on educational staffs, students, policymakers, administrators, parents, or schools). When used responsively, the UEEs yield intended positive effects regarding students' knowledge, skills mastery, performance, etc. On teaching side, for example, these tests help teachers to align their instruction with standards, and to develop their coaching practice professionally. On the part of the high-status stakeholders, the UEEs results interpretations and uses have potential positive effects on policymakers and administrators' ability to monitor and examine educational policies, and to promote any required changes or constructive alterations in the country's program accordingly.

In an opposite vein, however, the UEEs have seriously introduced negative consequences to various stakeholders involved at different layers of the program. Since the stakes are high in this country, especially in rationing future

opportunity as the basis for determining admission to the higher levels or to the employment opportunity as described above, the UEEs unintended consequences emerged to have overridden the negative ones. These negative consequences entailed harmful effects such as anxiety, opposing attitudes and despair at times a failure is caused, loss of motivation, etc. Apart from such psychological upshots abounding the learning atmospheres, it is clear that teaching and learning at schools became more test-like both in content and form of what is tested. As a result of targeting the higher scores in the UEEs, students' amount and quality of learning have been so much affected that they have tempted to overvalue the test-needed skills and technical efficiency to make test scores go up. Teachers alike are stimulated to adapt to complex, situation-specific circumstances altering their teaching to ensure improved test performances of learners since the UEEs results are used for competitive and accountability purposes. Such demands added greater pressures on teachers "to hurry along the curriculum in order to jettison good constructivist learning-, and learner-centered practices" (Glenwright, 2002).

All the above evidences clearly reveal that the main criticisms on sustaining with the UEEs in our nation's schools are "a narrowed curriculum", "reduced opportunities for deep learning", "reduced teaching time", "limited opportunity to assess higher order thinking skills", "no role for students' years of school background", and a number of other comparable challenges that all called for the new policy on university admission criteria.

2.2. The HNAEs as the UEEs Alternative

Generally speaking, what is targeted to be assessed by the high-stakes UEEs has strongly determined what gets taught and learnt at the lower stakes. This has a further consequence of hampering the learning-enhancement mission of the assessment, most occasionally. Consequently, unintended negative outcomes have been potentially triggered compared to those intended outcomes that have been essentially played out by the policymakers. Given this reality, what is at the stake is that reforming the existing examination system to overcome the possible negative upshots will enhance the attainment of learning objectives that underlie any assessment program. These debates finally led into the year 2007's Act with the

purpose of introducing major changes into the system of education such as bringing about positive effects on students' learning as well as overcoming deleterious effects of the high-stakes entrance examinations program.

As described before, the Act of 2007 postulates that assessing student cognitive outcomes through the HNAEs must gradually replace measuring the same outcomes through the sustained UEEs. The HNAE tests are annually provided by the Central Testing Office (CTO) of the Ministry of Education for almost nine courses annually passed by all third-grade students of all high schools across the country. The CTO institute determines the construction and rating process of such high school achievement tests so meticulously that the tests are not biased against certain subgroups or castes. Heading for the new assessment program, students of all four grades of high schools and pre-university schools are supposed to take the HNAEs on a four-successive-year progression (the Act of Student Selection, 2007: 4), rather than taking only the third-grade HNAEs currently available in the Ministry of Education.

According to policymakers and the

Parliaments' Members (personal communications, 2010-2011), in addition to the UEEs negative consequences discussed above, main reasons for the purposeful inclusion of the HNAE tests, are the essential problems with the current UEEs. These reasons mainly include "Multiple Choice-Format", "One-test-performance criteria" rather than "cumulative and formative school performances", "Norm-referencing rather than Criterion-referencing", and "Pure Accountability demands". As such, sustaining with the UEE program as a basis for making valid inferences about the educational performance runs counter to the professional principals asserting that a single test score should never make the basis for the important decisions for individuals. Compared with the UEEs, the known HNAEs are more aligned with the curricular and instructional proviso in the country to date, also include items on not only recognition but also on production skills too.

3. A Critical Appraisal: Looking through the HDM Evaluation Framework

Reform initiatives of the early 2000s imply that desired curricular objectives

can be put into high-stakes tests reform to shape and pull pedagogical and learning practices in desirable ways that will eventually improve instruction and learning quality. Such concerns with a host of others mentioned above formed the logic of the arguments made by a number of policymakers who hold the status that improved testing will trigger improved student performances and high-quality learning. However, experience suggests that the arguments of such a remediation are sometimes misguided (Chapman and Snyder, 2000) and success is not necessarily ensured. In fact, teaching and learning are multi-input phenomena that might not be easily changed by manipulation of single factors but, in reality, the existence of a multiplex of factors makes a complex constellation of problems ahead. Therefore, to gauge the amount of risks and threats involved in implementing the introduced alternative, to get assured whether it operates as it should, to exemplify its underlying messages and rationales, as well as to scrutinize the feasibility and possibility of such a change, the theoretical framework of HDM is adopted here. This framework of evaluation as raised by Henrichsen

(1989) and Wall & Horak (2008) has the potential to monitor the program from its earliest stage, and to disseminate results about understanding the intermediate factors that are at play for the success of a program like the new student admission program of Iran.

3.1. Antecedents Context

Taken for the purpose of the present evaluation panorama, Henrichsen's HDM initially postulates the newly set criterion (the HNAE) as '*a change*' in admission program of the universities and then proceeds to evaluate processes by which this new criterion is diffused and accepted by the program receivers. According to the HDM, final consequences of the HNAEs are determined not only by the factors that work during the time the HNAEs are introduced to the system, i.e., during "*the process phase*", but also by a constellation of factors floating in "*antecedent situation*" (i.e., the immediate context in which the HNAEs are introduced). The relevant antecedents in our case, as was partly touched in section 2 include the long-lasting UEEs, their content and structure, the traditional type of teaching and learning, student performances and

progresses, individual' psychological status, teacher methodology, and the measurement practices which already influence the amount and types of the wash-back in the current system.

A systematic study at this phase, called "baseline study" (Weir & Roberts, 1994, Fekete et al, 1999) should seriously analyze four types of antecedents as shown in Graph 1: 1) *the characteristics of the intended user system* (e.g. classroom factors, cultural factors, economic factors, education administration, geographical factors, political factors, school factors, and technology in school); 2) *the characteristics of the intended users* (e.g., attitude to classroom teaching, to exams, to new ideas, levels of education, abilities; personal factors; interest; goals); 3) *traditional pedagogical factor* (e.g. coverage of syllabus, content of teaching, methods of teaching, methods of classroom assessment); and finally 4) *the experience of previous reformers* (the outcomes of the earlier attempts to change).

A very telling criticism of the HNAEs-based criteria begins with an examination of these four major sources of information interpreted as four "evidential links"

(Messick, 1996). Much of the criticism is heavily on the paucity of any open forum on the characteristics of the intended 'users' and partly on the characteristics of the intended 'user system' Prior to full implementation of the new criteria, these two integrated antecedent factors should have been deeply come under a close scrutiny. The focus of the existing tracks of the baseline studies in the country, if done for such a purpose, however, has been mostly on traditional pedagogical factors as well as on few user system characteristics such as classroom factors, educational administration, and school factors. Except for very few studies done on the detriments of the conventional criteria of the UEEs (for example, Hajforoush, 2002; Keivanfar, 2002; Rejali & Kheradpazhoh, 2002) during the last decade, the outcomes of piloting studies including teachers', students', local experts', and parents' feedbacks on the new criteria were not reported nor analyzed systematically.

Therefore, to remove the shackles of negative wash-back of the UEEs, not only is it important to investigate the users' (teachers, students and UEEs candidates) factors during the process phase but also

even more vital to investigate them before introducing the change to the educational context. Teachers' and students' factors that seem most relevant to a responsive implementation and dissemination of the HNAEs benchmarks are 'motivation to learn about the new admission program', 'altered teaching/learning experience', 'teaching/learning autonomy', and 'confidence in teaching/learning'. Insights from these factors that reveal how likely it is that the HNAEs will successfully take hold in the intended context is very scarce for the time being.

In other words, although the defenders of the HNAEs-based assessment have criticized the critics, there does not appear to be a methodologically solid study showing that this new criterion per se can improve students' learning as well as teachers' pedagogy. To track these idealized outcomes as put into policies by policymakers, there needs to run studies investigating whether other related factors have been simultaneously altered too, and where. With reference to this importance, it should be mentioned that education program in Iran has emerged to be dominated by two other Ts (Teachers & Textbooks) in addition to the radical T

(Tests). Because 'Teacher-centeredness', and 'Textbook-centeredness' are also crucially influential, yet with different degrees, the feasibility of the tests-led reform will not have penetrating impact if the textbooks materials and practices the books are based on are not coordinately changed. While in some places a textbook functions as supplement to the teachers' instruction, in our country a textbook serves as the basis for much of the input the learners receive, or the content and practices they are exposed to in their class. Recent decentralization initiatives in Iran have not significantly reduced centralized control of school resources (textbook, technology, etc. ...) yet, and schools are still assigned to apply conventional curriculum content and textbooks sent to them; moreover, are asked to establish these prescribed texts as the main HNAEs sources.

Feeling recently expressed by some Iranian university experts and high school teachers is that as long as the textbook regime exists in the ME, practiced textbooks and prescribed conventional materials will encourage teachers to reallocate time to those specific contents that are likely to be tested by the HNAEs

but not those contents and tasks required to foster student meaningful learning. Thus, despite reinforcing transformation policies toward dynamic interactive instruction/learning, didactic-oriented textbooks together with didactic-oriented teachers will expose students to the decontextualized knowledge and memorization as before. The unavoidable roles of the three dominant Ts show that evaluation review and a constant revision to the existing resources are necessary at this stage but seemingly the ME has fallen behind in this respect until now. In summary, despite endeavors to unlock the potentials through changing the UEE program to the one with a claimed systematic contribution to the educational improvement, non-existence of antecedent studies should have sounded alarm bells among policy makers and planners.

The complexity is not limited only to such a historical studies but the UEE-change campaign must also deal with factors that comprise the process portion (Process Phase) of the change.

3.2. Process Phase

Intended consequences cannot arise unless the key factors are properly set on the

agenda at the '*process*' phase. HDM at this level of scrutiny entails understanding of the '*source and message of the change*', '*plans and strategies for carrying out the change*', '*characteristics of communication*', '*receivers' awareness and interest*', and '*inter-elemental factors of the users, user system and resource system*'. In light of the first factor, it is worth mentioning that the prerequisites for student admission is laid out in the Parliament' Laws (2007) and records, but there is no separate official document, to the best of our knowledge, to be delegated to local educational staffs and teachers, providing them with clear message about the philosophy behind the change or soliciting for any subsequent significant change in teaching, learning, and assessment contexts. In a series of interviews (Authors, 2011, a work in progress) done by the present researchers, more than 20 teachers and experts reported their incompetence about both the message of the present change and the plans and strategies that are employed to implement it by the officials.

Such an ambiguity and imprecision of the message(s), subsequently, have slowed down setting out specific plans for classrooms by change agents like teachers.

Teachers' preliminary knowledge and some of their ideas about the nature of their teaching plans or other plans required for implementing the new admission program, moreover, their coping with such a new program have not been empirically tapped yet. Above this level, when it comes to the implementation plans and strategies at higher organizational level, it has been observed that some of the main examination scenarios of the ME have been retarded too. Designing different HNAEs which can come along with the curricular objectives together with deliberations on the Standard tests are still matters of controversies that have formed one of the most formidable critiques on the ME.

Another area of controversy caused by the current testing reform is about the 'channel' of communication. If the channels of communication are not well chosen or the message gets distorted in the process of transmission, then there will be little chance that the receivers gain the awareness required for making reacts as the originators of the innovation intended (Wall and Horack, 2008). In case of Iran, policymakers advising on the new program hoped that the HNAEs would have positive

washback, or impact, on classroom teaching/learning and on the overall schooling program. The question is that whether the ME, MHE, SO and the Parliament communication channels about the new criteria are clear enough? Wall and Horak (2008) explained that Rogers and Shoemaker(1971).classified communication channels as either "mass media" or "interpersonal". The mass media channels in the current case includes the Parliament's web site and its archives for the passed laws, the ME's and MHE's documents, and SO's web site and weekly magazines. This seems to be working within the communication frame since 2007, however, the path from these authorized channels to the interpersonal ones (e.g. colleagues, students, school management, impact studies...) is not well-followed by the receivers. Although such authorized mass media are the most accessible sources at the time, local teachers, students, head-teachers, and parents mostly learn from interpersonal ones, though in differing degrees. Generally speaking, none of the interviewed teachers, heads, and experts (authors, 2011, a work in progress) mentioned any promotional leaflets from the central and local educational offices.

The controversy does not end here. Receivers' awareness, evaluation and making sense of the changed admission criteria is quite low at present: 2011. Of receivers, most high school teachers hopefully know about the nature of the HNAE tests and their major sections but they are not aware of the degree of contribution of the HNAEs to the final admission process, the weight of each course constituting the overall GPAs, the role and degrees of teachers' assessment in such a process, the types of assessment they can employ in their classrooms; nor are teachers aware of exams security and reliable scoring. Teachers need more exposure to the real and clear policies and practices of the new criteria to develop a good grasp of the change before giving their students clear messages about the standards that the change represents (Fullan, 1993). In addition to teachers, other receivers of the program, in particular, school managers, experts and students are better to heighten their awareness through official transmission of information rather than information of dubious quality through non-official sources. If receivers' intention to change towards the desired directions is an early

manifestation of their awareness and perceptions of the change, then the new criteria begin inducing its positive impact during the implementation phase.

Hence, in such a progression towards the intended impacts of the HNAEs, what remains to be explored by planners is that whether program receivers develop increasing awareness, understanding, and interest that are needed to ease their burden of planning and preparation to change.

4. Future: Consequences Facilitated or Hindered?

The UEEs reform is now under the way as a most widely favored assessment-led reform in Iran, paving the way for achievement of Macro-Concerns of educational assessment. The belief that this ideal can immediately solve a range of real educational problems rests on an often unarticulated theory, however. According to "Diffusion of innovation" premise (Rogers, 1983), a multitude of interplaying factors that contribute to the final accomplishment, sustainability, or rejection of a given change are at play. In-depth examination of these complex factors yields insights that will assist professionals and policymakers with planning and achieving the reform goals more efficiently. Though the ME has

recently announced interests in evaluating the impacts of the new criteria as one of its research priorities of the 2011 (Research Priorities of the ME, 2011), no methodical, systematic study on change impacts and the associated factors has been undertaken to date. The urgent need for such an evaluation is well argued by Brown (1995) who pointed out that evaluation is a heart that connects and gives blood to all other program elements. Evaluation in terms of change outcomes have been viewed so essential a process whose minimal ignorance will suffice to preserve the change as 'a still-born baby'.

Generally speaking, the UEEs reform program has been received positively as it appears to have given rise to a number of learning outcomes, nonetheless, there are points still left unaddressed in this new admission program for universities. As a significant area, for instance, there is a vital need to account for the way that the change is experienced by its end-users' perceptions since perceptions of a change, views on it, and attitudes toward it form the psychological part of the change outcomes. Markee (1990), Rogers (1983), and Stoller (1994) based on the conclusions of numerous empirical and non-empirical

studies cited a set of qualitative features that are perceived positively or negatively by the change agents and receivers. Based on their arguments, successful adaption of a change will most likely occur when the change is perceived to be highly feasible, explicit, relevant, and compatible, but not complex, inflexible, or uncertain. An analytic glance at the present testing-based reform reveals a list of similar factors that will, most likely, come facilitative or hindering in our context. Of these, the most important inherent features are 'complexity', 'compatibility', 'explicitness', and 'form'.

In brief, *complexity* is defined by Fullan (2001) as the "difficulty, skill required, and extent of alterations in beliefs, strategies, and use of materials" (p. 78). Successful diffusion of the new admission criteria requires knowing about the complexity involved at a number of levels, for example, at teacher, student, material, and logistic levels. In our case, regarding the skills required, for instance, both teachers and students have formal training, skills and relative qualifications for getting results from the HNAEs because they are familiar with the task types of the HNAEs. Yet, regarding the complex psychology of

perceptions and beliefs needed to take the HNAEs perspective into a serious account, the present change is not simple but seem rather complex.

In order to adopt and judge the acceptability degrees of the UEEs change, relevant to complexity perceived, '*compatibility*' feature should be also carefully analyzed. Compatibility acts as a cross-check on the perceptions that arise as a result of similarities versus differences observed in the two competing programs. A very subtle view is that "the HNAEs are exceedingly compatible to current practices of the ME and it is perceived that they likely facilitate the intended policies". Yet, surveys, interviews, and case-study results will indicate further empirically attested evidences on facilitating roles of the HNAEs compatibility.

As to the '*explicitness*' factor, potential users of a change should clearly get informed about the rationale, philosophy, specific goals and objectives boosting that change. Stoller (1994) puts an emphasis on the facilitative role of explicitness, visibility, and other semantically related terms through outlining the relevant literature on these features. In such vein, he argues for much likelihood of

developing favorable attitudes toward a given change when that change is perceived to be visible and explicit. In the present context, there are indications that lack of clarity of objectives and procedures, as well as invisibility of different outcomes of the HNAEs when completely adopted as a sustained admission criteria in future may hamper reform efforts and finally threaten those agents who might be satisfied with the status quo of student admission. Stakeholders' feedback on the explicitness of information about the new criteria has not been taken seriously yet; nor have sufficient feedbacks been obtained on the contribution of such a perceived explicitness to making a solid impression by the receivers.

Finally, of importance in future adoption is the '*form*' factor that not only supports the types and the nature of the activities and materials employed in classrooms and in preparation for the HNAEs exams but also gives a way to 'how to manage high school classes for future'. The merit is that ME's teachers are informed about the present HNAEs activities and materials but they seem not to have concrete ideas about how to plan

their classes in future if the HNAE-based program completely replaces the UEE-oriented program. Conceivably, Iranian teachers start introducing new activities into their classes only when the change has been translated into teaching materials. Accordingly, it will not be unusual that teachers will continuously assess their students based on the existing activities of the existing ME's books and of other materials which are not, for the time, in congruence with constructive theories behind the "assessment culture" (Segers & Dochy, 2001) but in reality are in consistency with the psychometric or "testing culture" (Wolf, Bixby, Glenn and Gardner, 1991). Such concerns about the 'form' of a change illustrate that 'base-line' evaluation is ubiquitous and has important implications for maintaining and improving the tests-based reform in Iran.

In nutshell, the currents of the future consequences and impacts perhaps resemble to the "wind and water" (Datta, 2009: 163) whose high-altitude currents flow in one direction and low-altitude on another. Put this in other words, in this *ocean metaphor* the surface layer (lower altitudes) and the deep layers (higher

altitudes) may moves in opposite ways. At present, the surface layers seems to favoring trials for finding out how effectively the new program and its underpinning policies promote student deep learning and attainment. However, the deeper layers (higher altitudes) include diversity of all traditions discussed above. It is through the powerful lens of the evaluation telescope that a fortuitous position for examining such diversity is achieved. "*Tradition has guided transition in Iran's curriculum-assessment reforms*" but unless the multi-dimensionality of a change of such an enormity is thoroughly evaluated from a multiplex perspective, there will be no guarantee that the intended practices follows the adopted policies.

References

- [1] Authors (2011). *British Journal of Educational Psychology*.
- [2] Alderson, J. C. (1986). Testing English for Specific Purposes: how specific can we get? *ELT Documents*, 127: 16-28.
- [3] Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-73.
- [4] Birenbaum, M. (2007). Evaluating the

- assessment: sources of evidence for quality assurance. *Studies in Educational Evaluation*, 33, 29-49.
- [5] Brown, J. D. (1995). *The elements of language curriculum: a systematic approach to program development*. New York: Heinle & Heinle.
- [6] Chapman, D. & Snyder, C. W. (2000). Can high stakes national testing improve instruction: reexamining conventional wisdom. *International Journal of Educational Development*, 20, 457-474.
- [7] Cheng, L. (1999). Changing assessment: Washback on teacher perceptions and actions. *Teaching and Teacher Education*, 15, 253-271.
- [8] Datta, L. (2009). Science-Based Educational Evaluation and Student Learning in a Global Society: A Critical Appraisal. In Ryan, K. & Cousines, J. B. (Eds.), *The Sage International Handbook of Educational Evaluation* (pp.149-168). U.S.A: Sage Publication.
- [9] Darling-Hammond, L., Ancess, J., & Falk, B. (2001). *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.
- [10] Decker, D. M. & Bolt, S. E. (2008). Challenges and Opportunities for Promoting Student Achievement through Large-Scale Assessment Results: Research, Reflections, and Future Directions. *Assessment for Effective Intervention*, 34 (1), 43-51.
- [11] Fekete, H. Major, E. & Nikolov, M. (1999). *English language education in Hungary: A baseline study*. Budapest: The British Council Hungary.
- [12] Fullan, M. G. (1993). Why Teachers Must Become Change Agents? *Educational Leadership*, Vol.50 (6).
- [13] Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). London: Cassell.
- [14] Glenwright, P. (2002). Language proficiency assessment for teachers: The effects of benchmarking on writing assessment in Hong Kong schools. *Assessing Writing*, 8 (2), 84-109.
- [15] Hajeforush, H. (2002). *Negative Consequences of Entrance Exams on Instructional Objectives and a Proposal for Removing them*. Evaluating the Issues of the Entrance Exams: Proceedings of the Esfahan University Conference on, Esfahan University, May.

- [16] Henrichsen, L. E. (1989). Diffusion of innovations in English language teaching: the ELEC effort in Japan, 1956-1968. New York: Greenwood Press.
- [17] Honig, B. (1987). *How assessment can best serve teaching and learning*. Assessment in the service of learning: Proceedings of the 1987 E¹S Invitational Conference. Princeton, NJ: Educational Testing Service.
- [18] *Iran's 20-Year Vision* (2005). The Council of Iran's Expediency. Retrieved Sep. 2009 from <http://www.majma.ir/Contents.aspx?p=67e04aa-7171-4f72-bdf7-e6f68c3547e5>.
- [19] Keivanfar, M. R (2002). Entrance Exams Consequencies and Inconsistency of their Items with the Quality of Education. Evaluating the Issues of the Entrance Exams: Proceedings of the Esfahan University Conference on, Esfahan University, May.
- [20] Kia, A. & Bozorgi, K. (2006). *Comments on "The Proposal for Student Admission to Higher Education Universities"*. (In Farsi). Technical Report for the Research Center of the Parliament, no. 8247.
- [21] Kikuchi & Browne (2009). English Educational Policy for High Schools in Japan : Ideals vs. Reality. *RELC*, 40, 172-191.
- [22] Kornhaber, (2004). Appropriate and Inappropriate Forms of Testing, Assessment, and Accountability. *Educational Policy*, 18 (1), 45-70.
- [23] *Letters to the RCP* (2003). Proposal for Changing University Entrance Examinations Program. The Isfahan House of Mathematics & RCP *Achieves*, 9-71/4879; 9-71/5184; 9-71/5478
- [24] Markee (1990). *The Diffusion of Communicative Innovations and Classroom Culture: An ethnographic Study*. University of Illinois at Urbana-Champaign.
- [25] Messick (1996). Validity and Washback in Language Testing. *Language Testing*.13, 251-256.
- [26] Munoz, A. P. & Alvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, 27 (1), 33-49.
- [27] Noble, A.J., & Smith, M.L. (1994b). *Old and New beliefs about measurement-driven reform: The more*

- things change, the more they stay the same. CSE Technical Report 373. Centre for Research on Evaluation, Standards, and Student Testing. Arizona State University.
- [28] Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- [29] *Philosophy of Education* (2009). Ministry of Education of the Islamic Republic of Iran. Retrieved August 2010 from <http://www.sce.ir/>.
- [30] *Research Priorities of the ME* (2012). Research Institute for Education, retrieved May 2011 from <http://rie.ir/index.aspx?siteid=82&pageid=1302>.
- [31] Rejali, A. & Kheradpazhoh, S. (2002). *Suggestions from Math Teachers on Negative Consequences of Entrance Exams*. Evaluating the Issues of the Entrance Exams: Proceedings of the Esfahan University Conference on, Esfahan University, May.
- [32] *Roadmap of the Official and General Education* (2009). Ministry of Education of the Islamic Republic of Iran. Retrieved August 2010 from <http://www.sce.ir/>.
- [33] Rogers, E. M. (1983). *The diffusion of innovations* (3rd ed.). London: Macmillian.
- [34] Ross (2008). Language testing in Asia: Evolution, innovation, and policy challenges. *Language Testing*, 25 (1), 5-13.
- [35] Segers, M., & Dochy, F. (2001). New assessment forms in problem-based learning: The value added of students' perspective. *Studies in Higher Education*, 26 (3), 327-343.
- [36] Shojaee, M. & Gholipoor, R. (2005). *Recommend Draft of Applying University Student System Survey and Designing Acceptance Model of University Student*. Research Center of the Parliament, No. 7624.
- [37] Stoller, F. (1994). The diffusion of innovations in intensive ESL programs. *Applied Linguistics*, 15 (30), 300-327.
- [38] Stoyhoff, S. (1989). Successfully implementing educational change and innovation. *ORTESOL Journal* X:17-33.
- [39] *The Act of Students Admission to Universities* (2007), (In Farsi). The Parliament of the Islamic Republic of Iran. Parliament Achieve, 132730.

- [40] *The Document of National Curriculum* (2009). Ministry of Education of the Islamic Republic of Iran. Ministry Archive.
- [41] *The Document of National Curriculum* (2010). Ministry of Education of the Islamic Republic of Iran. Ministry Archive.
- [42] Wall, D. & Horak, T. (2008). *The impact of changes in the TOFEL examination on teaching and learning in central and eastern Europe: phase 2, coping with change*. ETS, TOEFL.
- [43] Waters, M. (1995). *Globalization*. London: Rutledge.
- [44] Waters, A. and Vilches M. L. C. (2005). Managing innovation in language education: A course ELT change agents. *RELC Journal*, 36: 117–136.
- [45] Weir, C., & Roberts, J. (1994). *Evaluation in ELT*. Oxford, UK: Blackwell Publishing.
- [46] Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.
- [47] Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. In G. Grant (Ed.), *Review of Research in Education*, 17.

تحول در نظام پذیرش دانشجو به سبک کنکور سراسری در ایران: نقدی بر بستر تغییر

شایسته فر^۱، غلامرضا کیانی^۲، غفارثمر^۳، رامین اکبری^۴

تاریخ پذیرش: ۹۱/۱/۲۶

تاریخ دریافت: ۹۰/۷/۱۲

مطالعات نظری و تجربی گوناگونی که در کشورهای متفاوت در زمینه آزمونهای پر اهمیت سراسری انجام شده حاکی از تبعات غیر قابل انکاری است که با تغییر و تحول این آزمونها در کیفیت یادگیری، استانداردهای تدریس، و پاسخگویی معتبر به وجود می آید. در این راستا، سیاستگذاریهای ارزشیابی دهه پیش در ایران نیز تحول اساسی در نظام کنکور-محور پذیرش دانشجو را پذیرا بوده است. آنچه منجر به ایجاد بارقه این تغییر بسیار مهم گردید، در حقیقت، افزایش آسیب ها و آثار منفی مزمنی است که کنکور سراسری طی سالیان متمادی بر برنامه درسی، نحوه و نوع یادگیری، آموزش، و ارزشهای اجتماعی بر جای گذاشته است. اما این نکته را نباید دور از نظر نگه داشت که در بسیاری از موارد تغییر و تحول در آزمونها با موفقیت از پیش ترسیم شده و منظور نظر سیاستگذاران خاتمه نمی یابد. در رفع این معضل، پژوهش های موجود (ویر و رابرت، ۱۹۹۴؛ فکت و همکاران، ۱۹۹۹؛ وال و هوراک، ۲۰۰۸) به وضوح نشان از لزوم انجام مطالعات زمینه ای دارد، یعنی بررسی و ارزشیابی بستر تغییر و ویژگیهای زمینه ای که تغییر در آن آغاز و بسط می یابد. لذا بررسی میزان موفقیت، عملی، و مفید بودن تحول نظام پذیرش دانشجو در قالب چنین مطالعات زمینه ای دو چندان ضروری به نظر می رسد. از اینرو، نویسندگان این مقاله با اتخاذ مدل نظری هنریچسن (۱۹۸۹) یعنی "مدل ترکیبی بسط تغییر" که ارزشیابی هر تغییر آموزشی از چندین منظر متفاوت را لازم دانسته، ضمن توجه به ویژگی ها و عوامل موجود در بستر فعلی تغییر به ارزشیابی نقادانه تحول ایجاد شده در نظام پذیرش دانشجو می پردازند. در این مقاله ابتدا به سیاستگذاریهای علمی و همچنین محتوی مرتبط در اسناد رسمی بالادستی که زمینه ساز ایجاد تغییر و تحول در نظام ارزشیابی یادگیری به ویژه تحول آزمونهای سراسری و نظام پذیرش دانشجو بوده پرداخته می شود. سپس سیاست ها، مبانی نظری، و تبعات عملی آزمونهای کنکور سراسری در کنار آزمونهای پذیرش دانشجو در ملاک جدید (سابقه تحصیلی داوطلبان) به طور مبسوط ارائه می گردد. نتایج تحلیلی این تحقیق ضمن اینکه قابل تغییر بودن سیاست ها و پویائی رویکردهای ارزشیابی را مورد توجه قرار می دهد، به تبیین و تحلیل فرصت ها، تهدیدها، و چالش های پیش رو در تغییر و تحول اتخاذ شده می پردازد.

واژه گان کلیدی: آزمونهای پر اهمیت کشوری، تحول نظام ارزشیابی، پیامدها و اثرات آزمون، سابقه تحصیلی.

۱. دانشجوی دکتری گروه زبان انگلیسی دانشکده علوم انسانی، دانشگاه تربیت مدرس تهران.

۲. دانشیار گروه زبان انگلیسی دانشکده علوم انسانی، دانشگاه تربیت مدرس تهران.

۳. استادیار گروه زبان انگلیسی دانشکده علوم انسانی، دانشگاه تربیت مدرس تهران.

۴. استادیار گروه زبان انگلیسی دانشکده علوم انسانی، دانشگاه تربیت مدرس تهران.

