Interrelations between Policymakers' Intentions and School Agents' Interpretation of Accountability Policy in Israel

Lisa Amdur
Tel Aviv University
&
Irit Mero-Jaffe
Beit Berl Academic College
Israel

Abstract: The study examined the interrelations between policymakers' intentions for test-based accountability policy, and school agents' perceptions and actions with regard to this policy. Mixed-methods were used and encompassed 24 policymakers, 80 school principals, 168 teachers and case studies of four schools. New institutional theory, including the concept of "environmental shift" (Powell & Di Maggio, 1991) and the metaphor of "coupling" (Weick, 1976), served as the conceptual framework. Findings indicate that the interrelations between intentions, perceptions and actions are mainly tightly coupled. These are discussed by invoking three types of institutional isomorphism (DiMaggio & Powell, 1983): coercive, mimetic and normative.
Keywords: Accountability; loose coupling; public policy; policymakers; national testing

Interrelaciones entre las intenciones de los responsables políticos y las interpretaciones de la política de rendición de cuentas por agentes escolares en Israel
Resumen: El estudio examinó las interrelaciones entre las intenciones de las autoridades para la política de rendición de cuentas basada en pruebas y las percepciones y acciones de los agentes escolares con respecto a esta política. Métodos mixtos se utilizaron y abarcó a 24 políticos, 80 directores de escuela, 168 profesores y estudios de caso de cuatro escuelas. La nueva teoría institucional, incluyendo el concepto de "cambio ambiental" (Powell & Di Maggio, 1991) y la metáfora del "acoplamiento" (Weick, 1976) sirvieron de marco conceptual. Los hallazgos indican que las interrelaciones entre las intenciones, las percepciones y las acciones están estrechamente "acopladas". Éstos se discuten invocando tres tipos de isomorfismo institucional (DiMaggio & Powell, 1983): coercitivo, mimético y normativo.
Palabras clave: Rendición de cuentas; bajo acoplamiento; política pública; Políticos; Pruebas nacionales

Inter-relações entre as intenções dos decisores políticos e as interpretações da política de responsabilização dos agentes escolares em Israel
Resumo: O estudo examinou as inter-relações entre as intenções dos formuladores de políticas para a política de responsabilidade baseada em testes e as percepções e ações dos agentes escolares em relação a essa política. Métodos mistos foram utilizados e englobavam 24 formuladores de políticas, 80 diretores de escolas, 168 professores e estudos de caso de quatro escolas. A nova teoria institucional, incluindo o conceito de "mudança ambiental" (Powell & Di Maggio, 1991) e a metáfora do "acoplamento" (Weick, 1976), serviu de quadro conceitual. Os achados indicam que as inter-relações entre intenções, percepções e ações estão bem "acopladas". Estes são discutidos invocando três tipos de isomorfismo institucional (DiMaggio & Powell, 1983): coercivo, mimético e normativo.
Palavras-chave: Responsabilidade; acoplamento solto; políticas públicas; Formuladores de políticas; Teste nacional

Introduction

Within the literature on public policy, there is extensive research on both policy determination and policy implementation; however, there is less research with regard to the translation of policy from intention to practice. Within the field of educational policy, accountability, the process of evaluating schools based on student performance (Figlio & Loeb, 2011), is prominent in education systems across the world. Indeed, it is believe that accountability extends down into the system and increases the influence of political authorities in shaping classroom behavior (McDonnell, 2013). Thus, the purpose of the current study is to examine the relationship between the policy intentions of policymakers for accountability, and the interpretation of that policy by school agents. An understanding of this relationship may shed light on the extent to which policy implementation is consistent (or inconsistent) with policy aims (Spillane, Gomez & Mesler, 2009) with the primary focus being on the intentions of policymakers vis-à-vis the perceptions, interpretations and actions of school principals and teachers.

This study situates test-based accountability at the intersection between policy studies and institutional theory. The former views the determination and implementation of test-based accountability as a contextualized process involving two main agents: policymakers, who determine

policy, and school agents (principals and teachers) who implement it. The latter positions policy implementation within the broader context of complex organizational settings and relates to how implementers perceive and interpret policy, as well as what actions they take to address it (McDonnell, 1994a, 1994b). The theoretical framework utilized for an analysis of this relationship is new institutional theory (DiMaggio & Powell, 1983; Powell & DiMaggio, 2012).

Test-based Accountability

Accountability systems involve the collection of information regarding student performance, which is usually carried out through state-mandated tests, considered by some to be the "centerpiece" of accountability systems (Hanushek & Raymond, 2002; Hanushek & Woessmann, 2014). Policymakers, perceiving tests as being an effective policy instrument to influence education (McDonnell, 1994b), are attracted to them for a number of reasons. They perceive tests as being authoritative, objective, and fair (Shohamy, 2001; Shohamy, 2017). The visibility of tests provides an obvious indication of action (Linn, 2000) and thus policymakers are perceived to be promoting a sound educational agenda. Furthermore, tests are not only visible to the public, but the public also believes in tests.

Due to the relatively low cost, at least in comparison to other means of reform, testing is best suited to the constraints of government budget. Finally, tests can be implemented quickly which may be most significant to policymakers, considering their often short terms of office.

Intentions of Policymakers for Test-based Accountability

Policy intentions relate to the purpose and expectation of a policy (Shohamy, 2001). Intentions not only determine actual policy but may also shape policy adoption and implementation (Dunn, Jaafar, Earl, & Katz, 2013; May, 2012; Unger, 2008). Policy has even been considered as a "transformation of intentions" (Placier, Hager & Hull, 2005, p. 96). More often than not, a range of intentions drives a given policy and this appears to be the case for test-based accountability, for which policymakers have multiple intentions.

One of the most prominent intentions for test-based accountability is school improvement and the enhancement of school performance (Cohen, Moffitt, & Goldin, 2007; Malen & Rice, 2008). Another intention is to raise student achievement (Hamilton, Stecher, & Klein, 2002). When studying policy intentions through interviews with policymakers McDonnell (1994a), a leading scholar in the field of public policy, detected an additional intention for test-based accountability whereby policymakers use tests with the intent to heighten the awareness of teachers and administrators regarding instructional goals and expectations. Another intention has been found to be motivating teachers to work harder and make more effort (Firestone, 2014; Linn, 2004). Some scholars have argued that it is a means for top-down command and for wielding control (Brindley, 2008; Moe, 2003).

Policy Presentation

Policymakers are aware of the importance of gaining the attention of the implementing agents (Spillane, 2002). This is accomplished in many ways, primarily through laws, directives and regulations. However, the policy message may also be transmitted through other channels, for example through the provision of inducements (Bardach, 1977). Apart from gaining the attention of future implementers, policy presentation is designed to encourage implementing agents to think differently about their behavior in order for policies to work as levers for change and reform (Weiss, 1990). Thus the present study focuses not only on intentions but also on ways through which policymakers have presented their test-based accountability policy to school agents.

Policy Implementation

Policy implementation can be perceived as "the connection between the expression of governmental intention and actual results" (O'Toole, 1995, p. 43). Thus, implementation should correspond to intent and those who implement policy are required to interpret and understand policy and take action accordingly. However, implementation does not always play out according to the original blueprint set out by policymakers and there may be what Marshall (1998) has coined a "slippage between policy intent and outcomes" (p. 103).

Furthermore, policy implementation is a contextualized process, typically occurring in complex organizational settings wherein implementers occupy different levels of hierarchy (O'Toole, 2000). In the case of educational policy, its implementation depends specifically on principals and teachers (McLaughlin, 1991; Russell & Bray, 2013) who are working within the complexity of a school organization and who are expected to have the required capabilities (Cohen, Moffitt, & Goldin, 2007). It is often adapted to the school and classroom contexts in different ways than may have been intended (Hill, 2001).

Even though the responsibilities and professional demands of principals and teachers differ, both groups face the same policy expectations; both must interpret policy intentions and expectations (Heineke, Ryan & Tocci, 2015; Singh-Pillay & Alant, 2015) and must respond by setting a course of action. Thus, in understanding the implementation of test-based accountability it is necessary to focus on these individuals, specifically paying attention to their perceptions of policy and resultant actions as well as their educational context.

School Principals and Test-based Accountability

Studies have shown that school principals pay attention to state policy signals in general and focus on accountability policies in particular (Herman, 2007). When principals are confronted with the demands of test-based accountability, they make their own interpretations regarding the implications for them and their schools (Louis, Leithwood, Wahlstrom, & Anderson, 2010; Seashore & Robinson, 2012). In general, principals who perceive testing policy to be supportive of their work and valuable to the school are more inclined to support and implement it (Lyons & Algozzine, 2006).

The actions of principals may be divided into those on the level of the school and those on the level of the classroom. Actions on the level of the school include: planning resource allocation according to state testing demands (Dorn, 2007; Jones & Egley, 2004; Nichols & Berliner, 2007; Smith, Edelsky, Draper, Rottenberg, & Cherland, 1989), increasing oversight of teachers' work (Anagnostopoulos & Rutledge, 2007; Spillane, 2013), changing student classifications so as not to include them in the statistics (Figlio & Getzler, 2002; Glass 1990; McGill-Franzen & Allington 1993; Smith 1991), and analyzing test results and developing appropriate school plans (Herman, 2007).

Actions on the level of the classroom include supporting test preparation activities (Hamilton et al., 2007; Herman, 2007; Koretz, McCaffrey & Hamilton, 2001), helping teachers identify tested content, re-assigning teachers between tested and untested grades, presenting test results to teachers, providing support for curricular planning (Hamilton, 2003; Stecher & Chun, 2001), and providing professional development (Dorn, 2007). There are indications that principals tend to work harder in order to implement state policies (Firestone, Fitz & Broadfoot, 1999; Hill, 2001; Kelly, Odden, Milanowski, Heneman, 2000; Pedulla et al., 2003; Spillane, 2004).

Teachers and Test-based Accountability

Teachers perceive both the positive and negative aspects of test-based accountability (Louis, Febey, & Schroeder, 2005). Those who perceive the positive aspects of test-based accountability

believe that a system of checks for maintaining standards is necessary (Pedulla et al., 2003) and that there is value in testing as it ensures personnel will take responsibility for their actions (Robertson, 2003). Teachers perceive the tests as contributing to school effectiveness (Louis, et al., 2005), making them more aware of educational outcomes (DeBard & Kubow, 2002) and directing them to planned and focused instruction (Hamilton, 2003). On the other hand, teachers perceive test results as reflecting test-taking skills rather than language ability (Tahmasbi & Yamini, 2012). Furthermore many believe that test-based accountability decreases morale (Abrams, 2004), limits autonomy and harms professionalism (Crocco & Costigan, 2007; Leithwood, Steinbach & Jantzi, 2002; Rustique-Forrester, 2005).

With regard to teachers' actions, while some scholars have suggested that the influence of state-mandated testing is overstated and limited (e.g., Firestone, Mayrowetz, & Fairman, 1998; Grant, 2000), a much greater body of research provides evidence to the contrary, indicating both positive (Khattri & Sweet, 1996; Koretz & Barron, 1998), and negative aspects of implementation (Lipman, 2004; Nichols & Berliner, 2007; Watanabe, 2007). As teachers are the "final policy brokers" (Spillane, 1999, p. 144) their actions are critical in test-based accountability. In response to mandated tests, teachers re-allocate instructional time and work harder to cover more material effectively (Anagnostopoulos & Rutledge, 2007; Barrett, 2009; Firestone, et al., 1998; Firestone, Monfils, Schorr, Hicks, & Martinez, 2004; Jones & Egley, 2004). They focus more on achievement and extend the amount of time spent on test preparation (Hamilton & Stecher, 2006; Jones & Egley, 2004; Jones, Jones, & Hargrove, 2003; Koretz & Barron, 1998; Menken, 2008; Romberg, Zarinnia, & Williams, 1989; Shepard & Dougherty, 1991; Smith, 1991; Smith et al., 1989). Teachers suit the content taught to the test (Anagnostopoulos & Rutledge, 2007; Firestone et al., 1998; Jones & Egley, 2004; Koretz & Barron, 1998; Koretz, Linn, Dunbar, & Shepard, 1991) thereby narrowing the curriculum (Berliner, 2009; David, 2011), but do not necessarily make changes in pedagogy (Diamond, 2012). Finally when faced with such tests teachers have been found to collaborate more with each other (Grant, 2000; McLaughlin & Talbert, 2001). As with school principals, teachers have been found to work harder when faced with accountability demands (Firestone et al., 1999; Hill, 2001, Kelly et al., 2000; Pedulla et al., 2003; Spillane, 2004).

Theoretical Framework: "New Institutionalism"

The theoretical framework of new institutionalism has been used in a variety of organizational and institutional settings, not least being schools, as "public schools display all of the features addressed by institutional theory" (Mitchell, 1996, p. 174). This framework can provide the perspective through which to examine interactions between educational policies on the one hand, and school and classroom practices on the other (Burch, 2007, Burch & Spillane, 2005; Cerna, 2013; Coburn, 2004; Hasse & Krücken, 2014; Metz, 1990; Ogawa, 1992; Spillane & Burch, 2006).

Schools are considered to be interrelated organizations, which are open to external influence from their environment, or institutional field (DiMaggio & Powell, 1983). As such, schools must cope with diverse and sometimes contradictory demands and pressures filtering in from the institutional field. Of particular influence is an "environmental shift" that occurs when a component of the institutional field modifies an expectation or requirement of the organization within the field (Powell & DiMaggio, 1991; Rowan, 2002). One may consider state-mandated tests as influential environmental shifts that critically affect schools (Spillane & Burch, 2006).

New institutionalism has shown that schools generally seem to respond to environmental shifts in similar ways (Burch, 2007; Ogawa, 1992; Powell & DiMaggio, 1991; Rowan, 1982, 1995; Rowan & Miskel, 1999). For example, Hanson (2001) suggests that "the pressures on schools from

organizations and agencies in their environmental fields...are quite similar across the country, and in consequence, public schools in one region of the country tend to act like schools in other regions" (p. 648). Furthermore, school personnel develop their understanding of what is expected of them through participation in the same institutional environment (Meyer, Scott & Deal, 1981; Rusch, 2005).

Following from the idea that the institutional field exerts pressure on schools to which they respond, test-based accountability policy may exert pressure on the perceptions and actions of school agents (McLaughlin & Talbert, 1993; Rustique-Forrester, 2005; Spillane, 1999). These will be investigated in the context of a national test in Israel.

The Context

The Israeli education system is centralized and as such is characterized by the control of education by a central government body and a clear hierarchical structure. All public primary and secondary education, including school personnel, is government funded. Administrative and curricular decisions cascade down from the Ministry of Education and through the different levels of officialdom via reliance on inspectors who ensure the implementation of central regulations. Schools are required to follow national curricula that specify the standards and objectives to be pursued by all schools in all subjects and textbooks used in schools are to be approved by the Ministry. The Ministry also defines exit level achievement through a national matriculation exam (known as *Bagrut*).

In 2002, the Ministry of Education introduced a new, compulsory test battery, known as the Growth and Efficiency Measures (*GEMs*), into elementary and junior high schools. It was designed to provide information about three aspects of schooling: academic achievement in four core subjects – Hebrew/Arabic (as native language), Mathematics, Science and English as a foreign language (EFL); school climate; and pedagogical work environment. The information relevant to each of these aspects of schooling is collected through the following: achievement tests in grades 5 and 8; questionnaires completed by all students from fifth through to ninth grade; and interviews with school principals and teachers. The *GEMs* is designed and administered by the Ministry of Education. The Ministry also analyzes the results and provides individual reports to each school.

The school principal is responsible for making all the necessary arrangements for test administration, and for interpreting the findings and implementing any necessary changes based on the results. As for teachers, as it has been suggested that teachers of subjects tested respond differently to their institutional field than non-tested subjects (Spillane & Burch, 2006; Spillane & Hopkins, 2013) the focus of the present study is on EFL teachers. English is a compulsory subject taught in all public schools from fourth to 12th grade and as such has been included among the core subjects to be tested on the *GEMs* test battery.

The research questions studied within the context of test-based accountability (as exemplified in the *GEMs* tests) were:

1. What are the intentions of policymakers related to test-based accountability, and how are policy intentions presented to school agents (principals and teachers)?
2. How do school agents interpret the test-based accountability policy?
3. How can the interrelations between policy intention and implementation in the context of test-based accountability be characterized?

Method

A mixed-methods approach was chosen as it generates most useful results (Johnson, Onwuegbuzie, & Turner, 2007) particularly in accountability studies (Stecher & Borko, 2002) and within institutional theory (Weick, 1976). A qualitative approach was used to collect data from policymakers and school personnel as it provides more specific information and leads to deeper understanding (Firestone & Mayrowetz, 2000; LeCompte & Preissle, 1993). Explicit intentions appear clearly as written directives and articles of law, but alongside these are implicit intentions, which can only be tapped through interviews with the policymakers themselves. A case study approach allowed for a more in-depth study of the beliefs and actions of school agents, but it was supported by a quantitative approach, which allowed for the collection of information from a large number of participants and allowed for a description of general trends.

Participants

The participants included policymakers, school principals, EFL teachers and four schools.

Policymakers. A purposeful sampling technique served to select and recruit participants, with the intent of reaching all those policymakers who were directly or indirectly involved in the policymaking process (Creswell, 2007). This led to 18 policymakers who were holding or had held key government positions in education and were involved in setting testing policy in general and *GEMs* policy in particular. Six additional participants were recruited utilizing a snowball sampling technique (Warren & Karner, 2010) where we asked our initial 18 participants to connect us to any other policy actors who could serve as informants on the policy design. The final sample was thus comprised of 24 participants who reflected various levels of position in the hierarchy of the Ministry of Education, held roles with a range of responsibilities, and had participated in the preparation of policy documents and directives regarding testing policy.

The policymakers who participated in the study included three Director Generals of the Ministry of Education (past and present), the head of the Pedagogical Directorate who also serves as the Deputy Director General, the head of the Pedagogical Secretariat, two chief scientists (past), two subject chief inspectors, three district inspectors, heads of various divisions and departments in the Ministry of Education, including the Division for Elementary Education, the Division for Secondary Education, the Department of Evaluation, the Curriculum Division, the Department of Teacher Training, and District Offices.

Principals. Principals were randomly chosen from the official list of schools participating in the *GEMs* test during the year of data collection. Three-hundred principals were asked to complete a questionnaire of which 80 fully completed the questionnaire, representing a 27% response rate. Among the final sample of principals there was almost equal representation of female and male principals (56.2% female; 43.8% male); two-thirds administered elementary schools and one-third secondary schools. Half served in the secular sector, a quarter in the religious sector and a quarter in the Arab sector. Only a fifth of the principals had experience of three years or less and more than half had over nine years' experience.

EFL teachers. A purposive sampling procedure was employed by which questionnaires were distributed to EFL teachers attending compulsory training courses held at Centers for Teacher Professional Development. An effort was made to ensure that the participating teachers represented a valid sample of EFL teachers working in schools around the country. A total of 168 teachers participated in the study, the majority being female but with equal representation of elementary and junior high school teachers. Half of the teachers taught in the secular sector, a quarter in the

religious sector and a quarter in the Arab sector. Two-thirds of the teachers had over nine years teaching experience, and 80% of teachers held at least a bachelor degree.

Schools. The inclusion of a case study of schools was intended to provide additional information about the perceptions and actions of principals and teachers within their working context and thus enrich the data collected through quantitative means. Four schools participated in the case study. Those interviewed and observed within the case study of schools included all the school principals alongside administrative staff, and the EFL teachers in each of the schools. Schools were selected based on typical case sampling which is based on the rationale that these schools are illustrative as they are in no way extreme, deviant, exceptional or irregular (Patton, 2015).

Data Collection and Analysis

Data was collected over a one-year period, from spring 2005 to spring 2006 using a wide range of instruments. Data collection began with the collection of qualitative data from policymakers and continued throughout the year in the case studies of schools, through interviews and conversations with principals, administrative staff and EFL teachers. All the qualitative data was coded using preassigned theory-based and inductive data-driven codes (Bogdan & Biklen, 1998). All qualitative data was coded using Atlas.ti. The quantitative data was collected from principals and teachers during the school year through the use of specifically designed questionnaires.

Interviews. Three semi-structured interview protocols—for policymakers, school principals and EFL teachers—were developed and were based on the conceptual framework presented in the literature review. The semi-structured interview protocols included open-ended questions which were designed to allow for freedom of speech and encourage participants to pursue various directions of thought as they saw fit. The protocols underwent content validation and qualitative evaluation by four experts consisting of an expert in policy studies, an expert in assessment, an expert in educational administration and a teacher trainer. Interview questions were altered systematically as new information became available and themes emerged from on-going data analysis. Policymaker interviews included such topics as goals and priorities, views on testing, performance and achievement. Topics for school principals included impact of the *GEMs* prior to and following administration, performance expectations and consequences and general impressions of accountability system. Topics for EFL teachers included impact of the *GEMs* prior to and following administration, curriculum design, and classroom instruction and testing.

Observations. Observations of EFL teachers were carried out in a number of classrooms mainly, but not exclusively, in the grades participating in the *GEMs* (i.e. fifth and eighth grades). The main purpose of these observations was to complement the data collected through the interviews. The observations concentrated on the topics focused on by teachers in the classroom and on the types of materials that teachers used in the classroom, such as worksheets and practice tests used in test preparation. In addition to the observation of the classroom, observations of staff meetings were also carried out, particularly of those held prior to *GEMs* administration and upon receipt of the *GEMs* results. The foci of such observations were issues related to the *GEMs*, such as the division of labor in preparing materials for test preparation, and the discussion of test results.

Documents. Published documents related to official legislation were collected, such as the Director General Code of By-Law which is disseminated to all schools and available to the public through the Internet. In addition, the documentation of the official protocols of the Education Committee of the Knesset (Parliament) which are freely available on the Knesset Internet site was also collected. Documentation also included any written material prepared by the Department of Evaluation, such as PowerPoint presentations, meeting agendas, and written correspondence.

Questionnaires. Two questionnaires were developed to elicit information regarding the principals' and teachers' perceptions of and actions taken in response to the *GEMs*. The questionnaires were based on other surveys used in previous studies related to principals' and teachers' perceptions of state testing programs (Kennedy, Ball & McDiarmid, 1993) and on the literature. The questionnaires underwent content validation and qualitative evaluation by eight experts consisting of one national English inspector, two teacher trainers, an expert in language test design, an expert in educational evaluation, a school principal and two teachers.

Results

Test-based accountability is a prominent fixture in the centralized education system of Israel and one of its foremost policies. Its central feature, the state-mandated *GEMs* test administered to fifth and eighth graders in all public schools, has impacted those schools in various ways. The results of the current study relate to the *GEMs* test during its third year of implementation.

Intentions of Policymakers

The first research question related to policymakers' intentions for the test and its presentation. It was found that policymakers attach multiple intentions to the testing policy. One of the main intentions was to address the decline in student achievement as highlighted in the low ranking of Israeli students on international comparison studies. Many of the policymakers mentioned that "the beginnings of the problems were the disastrous results on the PEARLS and PISSA tests". Moreover, within a single year prior to the GEMs, three separate meetings held by the Education Committee of the Israeli Parliament were exclusively devoted to the topic. This problem was perceived as being particularly harsh as, in the words of one school principal, "the results are in the media and policymakers don't like bad publicity." A policy of additional testing to ensure success on future international tests seemed logical, especially as policymakers were aware of the advantages of testing. Moreover, there had been national tests in the past, thus policymakers saw the *GEMs* as a natural progression from previous state-mandated tests.

Intentions can be divided into those relevant to the level of the school and those relevant to the level of the state. With regard to the school, policymakers believed that "there was no accountability, kids go from grade to grade, whether they know the material or not, with no accountability." Thus, the *GEMs* was necessary in order to ensure school accountability including towards the state:

There must be accountability toward …the provider of the resources [the Ministry]. The bottom line is that the state gives money to the school principal so he can do something, not so he can waste the money, and he is required to report what he did with it.

Further, policymakers believed in the need to set minimum standards and common goals for all. "The number one goal of the *GEMs* is to reach certain standards, or agreed basic achievement that is clear to principals, teachers, and inspectors, what is the minimum expected from the pupil." The belief expressed by many policymakers who were interviewed was that, by getting schools not only to work toward common standards but testing schools to make sure that these common standards are achieved, there would be an automatic narrowing of the education gap. This sentiment is clearly expressed in the words of one policymaker: "The *GEMs* is a good infrastructure…it is an equalizing instrument that we can say all students need to reach."

A consistent finding and commonly cited intention of policymakers in the context of state-mandated tests is that the test ensured that certain subjects as well as specific content within a particular subject be taught and this is considered "the power of assessment…something that is assessed is of value." Another policymaker supported this by stating that "if there is no test in a

subject then it doesn't get taught…in our culture tests are a defining factor." With regard to defining content, a declared intention of policymakers was for EFL teachers to implement the national curriculum in their classrooms. It should be mentioned here that a new EFL curriculum had been introduced to schools just prior to the implementation of the *GEMs* and policymakers apparently thought to "take advantage of the *GEMs* as a trigger to get teachers working according to curriculum."

Besides the above intentions that related to the level of the school, policymakers mentioned two additional intentions that related to the level of the state. One was the intention of collecting data that would contribute to a more effective allocation of resources. The other was for evaluation purposes, though policymakers were aware that this may "be hard for principals to accept…that their work is checked".

Policymakers not only determine policy, but they are also aware of the need to gain the attention of policy implementers and motivate them to think about policy requirements. The findings indicate that policymakers did consider ways in which to present their policy to school agents: "You need to know how to market it correctly…so that the use will be right."

Policymakers presented policy by various means and through a range of channels. Policy was portrayed formally through government regulations and directives. Official documents included the Director General Code of By-Law which is binding for all public schools and the official test specifications. Informal communication was also harnessed as "sometimes when wanting to bring about change it is better to operate through channels that have the least resistance."

Informal communications were either written or oral, were more responsive to inquiries from the field than initiated by the policymakers, and were usually aimed at individual schools or school agents. Policy trickled down through the Ministry hierarchy thus: "The general inspector guides the district inspectors who guide the principals and the subject matter inspectors guide the subject matter coordinators in the schools." In-service training was another informal means of conveying policy. The Ministry organized "many conferences and study days for principals to learn about the *GEMs*". In the words of one policymaker "it's enough the school sees the syllabus of the in-service training given around the country to know what is expected of them."

Having discussed the policymakers' intentions for and presentation of policy, we turn now to the second research question and describe school agents' policy interpretation, considering their perceptions and their actions.

Perceptions of School Agents

School agents understood that the *GEMs* was "intended as a lever for improving achievement" believing that indications of the need for improvement came from the international tests. School principals were aware that one of the driving forces behind the *GEMs* was the low ranking of Israel on international studies: "if we go for the international standard then I believe that Israel won't want to be left behind, so they will have to have the *GEMs*." EFL teachers also seemed to be aware of this, as one EFL teacher succinctly put it "it's obvious those on high had to do something after everybody sees how bad we do in comparison to the rest of the world."

School principals and EFL teachers alike thought that one of the intentions was to ensure minimum standards for all students. School principals understood that the Ministry was interested in improving achievement and that the means to ensure it was by setting minimum standards. For one school principal it was clear that the Ministry "wanted to ensure a basic level that all the students in the country would reach…and by having the test the teachers would be oriented to this topic and they would work for the *GEMs*." Another school principal suggested that "there were some very weak populations that they [the policymakers] wanted to raise, so there was a need for a minimum

standard that everyone needs to reach." One vice-principal believed that "the *GEMs* is a lever for improving achievement…we saw that if we teach according to the *GEMs* we can improve achievement." From among the thirteen EFL teachers in the case study, five mentioned the idea of minimum standards. One EFL teacher believed that it was important that students "everywhere around the country" reach a certain level of English proficiency. This standard could be helpful in that "we're all on the same page." Another teacher believed that this insurance of a "standardized standard" would also ease the transition as "kids can go from junior high to high school without any problems."

School agents were aware of what was expected of them in terms of "action on the ground", specifically school principals were to prepare an action plan and teachers were to work according to the curriculum. In the words of one school principal: "I derive my action plan from the test, that was the intention and that is the use I make of the tests." Another principal added that "based on the test results I can set my school's aims and goals." The majority of principals who responded to the questionnaire understood they were expected to prepare an action plan. An EFL teacher explained that she understood policy intention to "make sure we [the teachers] are teaching the curriculum and we're teaching what we need to teach." Some EFL teachers thought that not only was the *GEMs* intended to ensure implementation of the curriculum but also when preparing tests they "should be based on the format of the *GEMs*."

Finally, some attested to understanding that policymakers were mainly interested in evaluating their performance: "I'm being measured as a principal, and my teachers are being measured as teachers." One junior high principal indicated "it [the *GEMs*] is to evaluate staff, but from a desire to treat staff favorably, not to entrap the staff, though some may believe that this is the case." However, one principal specifically stated that evaluation "serves as ammunition against the principals and teachers." Some thought that the school is evaluated in comparison with other schools as the "inspector compares schools." In the words of one EFL teacher "the test is per class, or per domain, that shows you it's testing the teachers."

Actions Taken by School Agents

With regard to the actions taken by school agents to address the *GEMs* it was found that school principals and EFL teachers took specific actions, mainly prior to the test in the context of test preparation, and their actions seemed to be synchronized and well-orchestrated. Even though EFL teachers were not as enamored with the test as school principals were, the actions of both school principals and EFL teachers were similarly directed toward preparing students for the test and ensuring their success on it. School principals and EFL teachers were less synchronized in relation to the use of test results.

School principals were found to support test implementation as they encouraged their staff to cooperate and follow all the directives related to test preparation and administration. Principals also paid particular attention to test preparation activities, allocating additional hours for test preparation and what some have called "*GEMs*-oriented learning", assigning the best teachers to the classes to be tested and ensuring that teachers could provide individualized attention to students in need. Their budget allocation also reflected the emphasis on test preparation as they provided the necessary funds for the purchase of booklets for test practice and the photocopying of test preparation materials and practice tests. In addition, principals advocated teacher training if and when they considered it necessary or appropriate. As test preparation and administration require extensive work and coordination principals sometimes delegated the administrative responsibility of the test to a pedagogical or assessment coordinator, who served as liaison between the school and

the Ministry, dealt with test administration procedures, monitored progress of test preparation, and was also responsible for the action plan subsequent to the receipt of test results.

Upon receipt of the *GEMs* report principals, together with the pedagogical or assessment coordinators, analyzed the strengths and weaknesses of their school, made changes where necessary, and an action plan was designed based on the data. In addition, they not only shared the results of the *GEMs* tests with the teachers, but also provided them with guidance in instructional planning and in identifying areas in need of improvement. Some school principals shared the results with parents, and some met with colleagues to compare results and discuss cooperative action. This was particularly evident between feeder elementary schools (in which the fifth grade tests take place) and their receiver junior high school (in which the eighth grade tests take place).

With regard to the actions of EFL teachers that address the *GEMs*, EFL teachers in both elementary and junior high schools focused on the test requirements and spent inordinate amounts of time exclusively devoted to test preparation. This was not only observed at the four schools participating in the case study, it was reaffirmed through the EFL teacher questionnaire. Based on the quantitative data 33% of the 134 participating EFL teachers reported devoting 11-20 hours to preparing their students for the test and a total of 21-30 hours of preparation time was reported by a quarter of the participant teachers. Taking into account that at the time of data collection the test was administered about eight weeks into the school year, and English lessons took place only three or four hours a week, one may conclude that most of the lessons prior to *GEMs* administration were almost exclusively devoted to test preparation.

In preparing students for the *GEMs* EFL teachers focused on the expected content based not only on the test specifications but also on published materials such as practice workbooks. EFL teachers also worked on promoting students' test-taking strategies, teaching the test and its format by giving practice tests made available by the Ministry on the Internet.

Data regarding ways in which teachers made use of the actual test results was limited, mainly due to the fact that the official report for each school only arrived toward the end of the year. Based on the existing data teachers made only limited use of the actual results included in the report that was sent to the schools as grades were reported on the level of the classroom and not on the level of the individual student. Thus, EFL teachers had an idea of the general trends in their classes, and if, for example, they perceived students to be weaker in writing they would include writing skills in their lesson planning.

Discussion

The discussion is based on the third research question with relates to the characterization of the interrelations between policy intention and policy implementation. Based on the findings, there is evidence that policymakers hold multiple intentions for the test-based accountability policy and school agents not only perceived these intentions, but also translated them into actions that were congruent with policymakers' expectations. In response to the test-based mandates, principals mainly believed that the test and accompanying report and action plan could bring order to an otherwise chaotic reality. Principals highlighted many of the positive aspects, including an emphasis on staff cooperation and better-planned budget allocation, as well as more attention to teacher professional training. Another positive outcome was a move toward more distributed leadership as principals, aware of the need to devote extensive time and energy to implementing the test, created a new role that involved responsibility for all aspects related to the GEMs. This role could also be expanded to include all aspects of assessment and testing.

Teachers, in response to test-based mandates and in line with the school policy as set out by their school administration, adapted their teaching. During the period leading up to the test teachers focused mainly on test requirements and on the expected content, and they devoted a lot of the available instruction time to test preparation. This sends a message to students that only succeeding on the test is important. Though teachers were found to be teaching skills that had previously not been taught, this entailed narrowing their syllabus, which meant that other content might have been neglected. Furthermore, to some extent teachers' academic freedom was reduced as content was dictated through the test.

It would seem that the principals in general have a more positive attitude toward the test than the teachers. This may be explained by the fact that the principals have more direct contact with the policymakers and with the channels through which the policy is conveyed. In this way they may have a better understanding of policy intentions. This also provides them with more of an opportunity for interaction with the policymaker in cases that they wish to clarify any of their misunderstandings. In addition, teachers are the ones who bear most of the burden in following the directives of the test and the success of their students on the test is a direct reflection on their performance. These pressures may explain the more negative attitude of the teachers toward the test. Though this study focused on a different education system, one that is centralized and has unique characteristics, these outcomes are not dissimilar to those found in many other educational contexts operating under test-based accountability.

Following the alignment of policy implementation with policy intention, institutional theory is invoked to characterize the interrelations between policy intention and implementation in the context of test-based accountability. Institutional theory offers "important insights for understanding contemporary [educational] policy and practice" (Burch, 2007, p. 93). Within the tradition of new institutional theory, the notion of loose coupling has been applied within organization studies for many years (Scott, 2004) and has been "especially prominent in explorations of educational organizations" (Swanson, 2005, p. 245). Loose coupling describes the match between different elements of an organization and elements in its institutional environment, such as governing agencies, or between the elements within the organization itself (Orton & Weick, 1990; Weick, 1976). Coupling elements refer to anything that may be tied or linked together (Weick, 1976). In the current study the coupling elements are the policymakers' intentions and school agents' perceptions and actions with regard those intentions.

The relations between schools and their environment have traditionally been considered loosely coupled (Weick, 1976). However, Spillane and Burch (2006) have argued that: "the ever increasing use of testing and systems designed to hold teachers and administrators directly accountable for student learning raise doubts about the assumption that schools...are loosely coupled systems" (p. 88). Diamond (2007) supports this assertion based on his study of schools in Chicago's accountability system where he found that high stakes testing policies are tightly coupled with classroom practice. Fusarelli (2002) has confirmed this and argued that due to test-based accountability "the tightness of coupling is far greater than at any other time" (p. 571). In the words of McDermott (2007), "the tightness takes the form of statewide standards, with mandatory tests…and accountability for results" (p. 81).

In line with these scholars, the findings in the current study point to tighter coupling between schools and their institutional environment. The coupling of intentions, perceptions and actions will now be discussed through four dimensions of coupling that characterize the quality of the relationship between coupling elements (Weick, 1982) – dependency, directness, strength and consistency. *Dependency* refers to the extent of dependence between coupling elements whereby loosely coupled elements are relatively autonomous of each other (Pfeffer & Salancik, 1978; Weick,

1976), and interdependence is characteristic of tight coupling. Interdependency was found between schools and government. School agents, by the nature of the centralized education system, are almost completely dependent on the government for their functioning, including funding, teacher allocation, and curriculum. It would appear that the extent of dependence between government and school is amplified within the context of the *GEMs* especially in the eyes of policymakers who believe that as they fund schools they have the right to test and monitor schools. Furthermore, the Ministry defined test content and conditions for test administration and schools were obliged to cooperate. This and more, the Ministry provided schools with test results and analyses. In addition, the Ministry approved a school's action plan prepared by school agents and based on test results.

Directness refers to how directly the coupling elements are coupled; schools are loosely coupled with the environment when school personnel communicate with officials through several levels of hierarchy rather than directly (Beekun & Glick, 2001; Weick, 1982) and conversely the more direct the linkages between elements, the tighter the coupling (Weick, 1982). Within the context of the *GEMs,* policymakers were more directly in touch with school agents, through letter writing, responsiveness to individual school agents who initiate contact, participation in study days and even school visits. This type of direct contact was not customary outside the context of the *GEMs.* Thus, it would seem that for the dimension of directness policymakers and school agents were tightly coupled.

Strength refers to both the frequency and intensity of contact between the coupling elements, and ranges from frequent and intense contact to infrequent and weak contact (Weick, 1982). The stronger the linkage between elements the tighter is the coupling (Weick, 1982). Besides being direct, contact between policymakers and school agents was also frequent and intense. This was even expressly mentioned when school principals were asked to compare a year with the *GEMs* to a year without and they decidedly mentioned the frequent contact between them and the Ministry. Training, study days and meetings all indicated the strength of contact and provide evidence for the tightly coupled relations between policymakers and school agents.

Consistency refers to the variation in response of the coupling elements to a given stimulus (Orton & Weick, 1990; Weick, 1982). It is evident that on this dimension too there was tight coupling between policymakers and school agents. Their responses within the context of test-based accountability were limited to being test-specific and goal-specific. With regard to being test-specific, the focus of interactions between policymakers and principals was restricted mainly to issues related to the *GEMs* test, specifically how to implement it, what is expected to be gained from it and making use of the results. With regard to being goal-specific, policymakers focused solely on the goals set by schools and the measures of success as detailed in a school's action plan. Similarly, meetings held between school agents and officials from the municipality were devoted to discussions related solely to the *GEMs.*

Taken together these four dimensions of coupling provide evidence of the tight coupling between policy and implementation. This begs the question of whether the tight coupling brings about school improvement, being the main intention of policymakers, or whether schools are only "going through the motions" so to speak. It would seem that some changes were not simply superficial or cosmetic. The tight coupling meant that directives and regulations became embedded in formal structures of schools as they adopted routines of alignment with the curricular content of the test and of monitoring of instruction. The creation of the administrative position of assessment coordinator provides further indication that government regulation had been incorporated into the formal structure of school (Spillane, Parise & Sherer, 2011). Some implications of this include greater distribution of leadership, changes in school assessment policy, and greater focus on

methods of testing and assessment, and analysis and interpretation of assessment results, to name just a few.

The tight coupling also meant the promotion of a stronger commitment to professional development. Both principals and teachers were required to participate in various workshops and study days offered by the Ministry. These opportunities for professional developed contributed to school agents' enhanced competences. The participation in common professional development also meant the creation of a "common language", something that in principle could bring about mutual understanding at all levels of the education system and between administration and staff in the schools.

The tight coupling evidenced between policy, administration and teaching, and the resultant conforming of schools to government mandates, may be explained by the process of institutional isomorphism (DiMaggio & Powell, 1983), which, as Fusarelli (2002) has argued, "serves to tighten the loose coupling of educational organizations" (p. 568). Indeed, institutional isomorphism has been used by institutional theory to explain why schools generally seem to respond to environmental shifts in similar ways (Burch, 2007; Ogawa, 1992; Powell & DiMaggio, 1991; Rowan & Miskel, 1999; Scott, 1991).

Isomorphism is defined as a "constraining process that forces one unit in a population to resemble other units that face the same set of environmental conditions" (DiMaggio & Powell, 1983, p. 66), in other words, that involves pressures towards conformity or compliance. Isomorphism can "cause whole populations and communities of organizations to act in concert as a result of their shared fate" (Rowan 2002, p. 61) and this seemed to be the case for the school agents in the current study.

Three types of pressures leading to institutional isomorphism (DiMaggio & Powell, 1983; Meyer & Rowan, 1977) include: *coercive* pressures, often involving political or government pressures and the force of the state to meet expectations and requirements, and including regulatory oversight, incentives, and control; *mimetic* pressures that ensure success in an uncertain environment by organizations drawing on habitual responses that had previously succeeded, by imitating or emulating other organizations' activities and successful peers, or by copying successful approaches of other organizations; *normative* pressures based on what is thought to be professional knowledge and how professionals should be or that rely on socialization (through professional organizations and professional media), which leads to viewing certain structures and processes as legitimate.

Coercive isomorphism was evident in the context of the *GEMs* as the test involved governmental regulation and political pressure exerted by the Ministry on the schools. Policymakers issued governmental directives, exercised control over the mechanisms of test development, administration and analysis and monitored the actions of schools within the context of the test. The actions of school agents reflected their compliance with these policy expectations. However, one may claim that compliance is inherent in a centralized education system, as by definition policy is top-down. Following the guidelines for the test may seem natural in schools that this has been there modus vivendi for all government initiatives.

Mimetic isomorphism was evident as schools tended to compare not only their test results but also the actions they had taken in order to achieve those results. They also turned to practices that had been successful in the past as similar tests had been given before. The sharing across schools may seem a positive action as it may contribute to shared understandings, learning from each other, and contribute to easing students' transition from elementary to junior high schools. However, this sharing could be derived not from a desire to share knowledge and expertise, or a learning community, but rather from competition. Whether this type of competition contributes to school improvement or not is an issue to be further pursued.

Normative isomorphism was evident as the test was considered the norm for school work and for judging school performance. The test was considered by school agents to be a legitimate means of collecting information about the status of education in general and for decision making within the school in particular. The assessment imparts legitimacy to the school; therefore, the test becomes the norm, or at least the normative way of evaluating schools and their performance, to the exclusion or omission of alternative means of assessment that may be more conducive to learning. Norms tend to be the most embedded within school culture, and thus they are not easily changed. Because the test becomes the norm, the school agents may be directed toward it to the exclusion of other, perhaps more authentic and valid forms of assessment.

Viewing test-based accountability through the prism of new institutional theory and harnessing the concept of loose coupling is helpful in explaining the response of different schools to accountability pressures (Burch, 2007; Ogawa, 1992; Powell & DiMaggio, 1991). In the context of the Israeli *GEMs* test it appears that the tightening of relations between policymakers and school agents has heightened the centralization of an already centralized system. The policy of accountability and the policy instrument of a state-mandated test are powerful regulatory mechanisms that have successfully brought about tighter coupling between the school and its institutional environment and it would appear that breaking away may be almost impossible.

Indeed, some have wondered why teachers have not become more socially responsible about the use of tests (Shohamy, 2001, 2007). It would seem that the pressures of the institutional environment may be so great that teachers "are bound to uphold the 'way things are' because other ways of thinking or acting have become largely unthinkable" (Stryker, 1994, p. 867). Scott (2001), a leading organizational theorist, concurs and believes that compliance is a result of a belief that, under certain circumstances, "other types of behavior are inconceivable" (p. 134).

This study has not presented an exhaustive range of responses to the test-based accountability system in place in Israel. However, the results do provide insight into ways in which the policy context of state-mandated tests affects the perceptions and actions of principals and teachers. Studies on policy implementation within accountability systems have highlighted patterns of tight and loose coupling between policy, school and classroom, and have shown that teachers and administrators are attending to policies and working hard to implement them (Firestone et al., 1999; Hill, 2001). The tightly coupled relationship between the policy environment and schools has resulted in schools focusing on the tests as an indicator of school quality (Mintrop & Trujillo, 2007, p. 319). The tight coupling has also meant that policy is exerting control over the curriculum (Au, 2007). This tight coupling between schools and their governing agencies has been confirmed in the current study and it is not surprising when one considers that the system studied is a centralized one and that the general intention of test-based accountability is to increase external control over what happens in schools and classrooms (Fusarelli, 2002).

Policymakers in the centralized education system, focusing mainly on school improvement, should be aware of both the positive and negative aspects of the different types of isomorphism, minimizing coercive isomorphism and highlighting the positive aspects of mimetic and normative isomorphism. This has the potential to reduce the pressure brought to bear by themselves on schools and allow for more school autonomy. It could also have a more positive effect on school agents' morale (Ladd, 2017). In keeping with normative isomorphism, policymakers may also wish to explore other avenues of action, such as allowing schools to implement their own accountability system (Polikoff, McEachin, Wrabel, & Duque, 2015). In such a case, policymakers concerned with school progress, could design a range of tools for monitoring schools. This is especially relevant for policymakers who may have too narrow a view of schooling (Brighouse, et al., 2016). Another suggestion that has arisen from the literature is moving away from achievement gaps across the state

to a focus on reducing within-school or within-district gaps (Ladd (2017; Neal & Schanzenbach, 2010). This would divert the focus from low-performing subgroups to reducing the gap within a school or district. It would also encourage teachers to make better use of assessment data, making the information gleaned from the test more meaningful and relevant. Focusing on the district rather than the state is particularly feasible in Israel as the system is divided into a number of districts and schools in the same district could pay more attention to all their students rather than pointing out a selected few.

It is necessary for school agents to not only understand policy messages and translate them into action; they should be aware of the possible sources of pressures on their work and interpret policy in a way that is not simply a direct response to such pressures (Dulude, Spillane & Dumay, 2017). School principals, being more aware of institutional pressures, may enhance the leadership skills necessary when faced with environmental shifts and their possible impact on schools. Programs designed for training future principals and teachers may benefit from insights gained in this study as to how school agents respond to policy mandates. Teachers may also be more empowered as they, too, become more aware of environmental shifts and their impact on their classroom practice. Pressures that seem to them as being unwarranted may be handled in ways that empower them rather than causing them to act in contradiction to their beliefs.

In the current study there was no attempt to analyze causative relations. However, there is a feeling that the test actually led to the response that was exhibited by schools both on the administrative level and on the level of the classroom. Of course, there may be any number of intervening variables that have contributed to the response of school agents to the policy. It is suggested that future research provide more insight into the causal relations between policy and policy implementation, perhaps through the design of a longitudinal study that closely follows schools over an extended period. A longitudinal design could also capture the long term effects of the policy. The local and more immediate responses of school personnel to accountability mandates, such as narrowing the curriculum, teaching to the test, and changing resource allocation, are clear. Indeed teachers and principals attest to the fact that they "get back to normal" after the test. The question remains to what extent the test has more far reaching consequences for the school.

This study focused on the main policymakers within the Ministry of Education as well as a number of officials external to the Ministry of Education, such as those in municipalities. However, other stakeholders holding influential positions may also be a source of important information, for example, officials from teachers' unions. It is recommended for future studies that officials from these unions be included as their response may provide more insight into the position of teacher representatives with regard to the accountability policy. It could also provide insight into the support teachers may expect if they were to consider a more activist stance as it is these unions that protect the rights of teachers. With regard teachers, the study focused on EFL teachers faced with the dictates of the *GEMs*. It is suggested that future studies focus on additional groups of teachers who are also faced with the same dictates, such as Math or Science teachers.

Another recommendation for future research is to focus on other sets of couplings. While a focus on couplings between educational organizations and their governing agencies was the center of the current study, there are additional coupling elements that are worth investigating as they are related to the effective functioning of schools. Under this heading would be included, for example, the relationship between teachers and their disciplinary societies or the coupling of schools with publishing houses and testing agencies. These issues of coupling are important not only to understand the work performed in schools as organizations in a given institutional field, but also to shed light on prospects for changing the nature of that work (Coburn, Hill, & Spillane, 2016).

In conclusion, a state-mandated test, conceived as an environmental shift through which isomorphic processes are invoked, has the power to bring about tight coupling between policymakers and school agents, between policy intention and policy implementation. This may be attributed to the joint and concerted effort on the part of the policymakers to ensure that schools comply with the accountability requirements. It may also be attributed to the fact that schools are strongly and directly dependent on governmental agencies. Whatever the case, school agents were aware of most of the intentions of policymakers for the test, and pragmatically translated policy into action, particularly in terms of preparing an action plan and following a national curriculum. Perhaps an understanding of policy intentions can temper the negative effects of top-down policy initiatives. Intentions being more transparent and implementers having a better understanding of policy intentions may allow for more mindful implementation of policy that is accompanied by a deep understanding of purpose and objectives. Having a better understanding of test-based accountability in the context of institutional theory could also contribute to school improvement. Taken together these could bring about more meaningful and lasting change.

Acknowledgements

References

Abrams, L. M. (2004). *Teachers' views of high-stakes testing: Implications for the classroom.* Education Policy Research Unit (EPRU), Arizona State University.

Anagnostopoulos, D., & Rutledge, S. A. (2007). Making sense of school sanctioning policies in urban high schools: Charting the depth and drift of school and classroom change. *Teachers College Record, 109*, 1261-1302.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*, 258-267. https://doi.org/10.3102/0013189X07306523

Bardach, E. (1977). *The implementation game: What happens after a bill becomes a law.* Cambridge: MIT Press.

Barrett, B. (2009). No child left behind and the assault on teachers. *Teaching and Teacher Education, 25,* 1018-1025. https://doi.org/10.1016/j.tate.2009.03.021

Beekun, R. I., & Glick, W. H. (2001). Organization structure from a loose coupling perspective: A multidimensional approach. *Decisions Sciences, 32*, 227-250. https://doi.org/10.1111/j.1540-5915.2001.tb00959.x

Berliner, D. C. (2009). MCLB (Much curriculum left behind): A U.S. calamity in the making. *The Educational Forum, 73*, 284-296. https://doi.org/10.1080/00131720903166788

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research in education: An introduction to theory and methods* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Brighouse, M. H., Ladd, H. F., Loeb, S., & Swift, A. (2016). Educational goods and values: A framework for decision makers. *Theory and Research in Education, 14*, 3-25. https://doi.org/10.1177/1477878515620887

Brindley, G. (2008). Educational reform and language testing. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education (2nd Edition). Language testing and assessment* (pp. 349-364). New York, NY: Springer Science. https://doi.org/10.1007/978-0-387-30424-3_187

Burch, P. (2007). Educational policy and practice from the perspective of institutional theory: Crafting a wider lens. *Educational Researcher, 36*, 84-95. https://doi.org/10.3102/0013189X07299792

Burch, P., & Spillane, J. P. (2005). How subjects matter in district office practice: Instructionally relevant policy in urban school district redesign. *Journal of Educational Change*, *6*(1), 51-76. https://doi.org/10.1007/s10833-004-7781-5

Cerna, L. (2013). The nature of policy change and implementation: a review of different theoretical approaches. Paris, France: OECD/CERI. Retrieved from https://pdfs.semanticscholar.org/15d8/a529d84b491b79d169aeabbfcb0532200971.pdf

Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education, 77*, 211–244. https://doi.org/10.1177/003804070407700302

Coburn, C. E., Hill, H. C., & Spillane, J. P. (2016). Alignment and accountability in policy design and implementation: The Common Core State Standards and implementation research. *Educational Researcher, 45*, 243-251. https://doi.org/10.3102/0013189X16651080

Cohen, D. K., Moffitt, S. L., & Goldin, S. (2007). Policy and practice: The dilemma. *American Journal of Education, 113*, 515-548. https://doi.org/10.1086/518487

Creswell, J. W. (2007). *Qualitative inquiry and research design*. Thousand Oaks, CA: SAGE.

Crocco, M. S., & Costigan, A. T. (2007). The narrowing of curriculum and pedagogy in the age of accountability. *Urban Education, 42*, 512-535. https://doi.org/10.1177/0042085907304964

David, J. (2011). What students need to learn: High-stakes testing narrows the curriculum. *Educational Leadership, 68*(6), 78-80.

DeBard, R., & Kubow, P. K. (2002). From compliance to commitment: The need for constituent discourse in implementing testing policy. *Education Policy, 16*, 387-405. https://doi.org/10.1177/08904802016003002

Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education, 80*, 285-313. https://doi.org/10.1177/003804070708000401

Diamond, J. B. (2012). Accountability policy, school organization, and classroom practice: Partial recoupling and educational opportunity. *Education and Urban Society, 44*, 151-182. https://doi.org/10.1177/0013124511431569

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48*, 63-82. https://doi.org/10.2307/2095101

Dorn, S. (2007*). Accountability Frankenstein: Understanding and taming the monster*. Charlotte, NC: Information Age.

Dulude, E., Spillane, J. P., & Dumay, X. (2017). High stakes policy and mandated curriculum: A rhetorical argumentation analysis to explore the social processes that shape school leaders' and teachers' strategic responses. *Educational Policy*, *31*, 364-403. https://doi.org/10.1177/0895904815598396

Dunn, R., Jaafar, S. B., Earl, L., & Katz, S. (2013). Towards data-informed decisions: from ministry policy to school practice. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education* (pp. 155-175). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-007-4816-3_9

Figlio, D. N. & Getzler, L. S. (2002). *Accountability, ability, and disability: Gaming the system?* Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w9307

Figlio, D. N., & Loeb, S. (2011). School accountability. In E. A. Hanushek, S. Machin & L. Woessmann (Eds.), *Handbooks in economics: Economics of education* (Vol. 3) (pp. 383-421). North-Holland, the Netherlands: Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00008-9

Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher, 43*, 100-107. https://doi.org/10.3102/0013189X14521864

Firestone, W. A., Fitz, J., & Broadfoot, P. (1999). Power, learning, and legitimation: Assessment implementation across levels in the United States and the United Kingdom. *American Educational Research Journal, 36*, 759-793. https://doi.org/10.3102/00028312036004759

Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change the effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis, 20*, 95-113. https://doi.org/10.3102/01623737020002095

Firestone, W. A., Monfils, L. F., Schorr, R. Y., Hicks, J. E., & Martinez, M. C. (2004). Pressure and support. In W. Firestone, L. Monfils & R. Schorr (Eds.), *The ambiguities of teaching to the test. Standards, assessment and educational reform* (pp. 63-90). Mahwah, NJ: Lawrence Erlbaum Associates.

Fusarelli, L..D. (2002). Tightly coupled policy in loosely coupled systems: Institutional capacity and organizational change. *Journal of Educational Administration, 40*, 561-575. https://doi.org/10.1108/09578230210446045

Glass, G. (1990). Using student test scores to evaluate teachers. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation* (pp. 229-40). Newbury Park, CA: SAGE. https://doi.org/10.4135/9781412986250.n14

Grant, S. G. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York state testing program. *Education Policy Analysis Archives, 8*(14). https://doi.org/10.14507/epaa.v8n14.2000

Hamilton, L. S. (2003). Assessment as a policy too. *Review of Research in Education, 27*, 25-68. https://doi.org/10.3102/0091732X027001025

Hamilton, L. S., & Stecher, B. M. (2006, April). *Measuring educators' responses to high-stakes testing.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (Eds.) (2002). *Making sense of test-based accountability in education.* Washington DC: Rand.

Hamilton, L. S., Stecher, B. M., Marsh, J. A., Sloan McComb, J., Robyn, A., Lin Russell, J…Barney, H. (2007). *Standards-based accountability under no child left behind: Experiences of teachers and administrators in three states.* Santa Monica, CA: RAND Corporation, MG- 589-NSF.

Hanson, M. (2001). Institutional theory and educational change. *Educational Administration Quarterly, 37*, 637-661. https://doi.org/10.1177/00131610121969451

Hanushek, E. A., & Raymond, M. E. (2002, June). *Lessons about the design of state accountability systems.* Paper prepared for Taking Account of Accountability: Assessing Policy and Politics Harvard University.

Hanushek, E. A., & Woessmann, L. (2014). Institutional structures of the education system and student achievement: A review of cross-country economic research. In R. Strietholt, W. Bos, J. E. Gustafsson, & M. Rosén (Eds.)*, Educational policy evaluation through international comparative assessment* (pp. 145-175). New York NY: Waxmann Verlag.

Hasse, R., & Krücken, G. (2014). Coupling and decoupling in education. In B. Holzer & T. Werron (Eds.), *From globalization to world society*, (pp. 197-214). New York: Routledge.

Heineke, A. J., Ryan, A. M., & Tocci, C. (2015). Teaching, learning, and leading preparing teachers as educational policy actors. *Journal of Teacher Education, 66*(4), 382-394. https://doi.org/10.1177/0022487115592031

Herman, J. L. (2007). *Accountability and assessment: Is public interest in K-12 education being served?* (CRESST Report 728). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Hill, H. C. (2001). Policy is not enough: Language and the interpretation of state standards. *American Educational Research Journal, 38*, 289–318. https://doi.org/10.3102/00028312038002289

Johnson, R. B., Onwuegbuzie, A. J., & Turner L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*, 112-133. https://doi.org/10.1177/1558689806298224

Jones, B. D., & Egley, R. J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education Policy Analysis Archives, 12*(39). https://doi.org/10.14507/epaa.v12n39.2004

Jones, M. G., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high-stakes testing.* Lanham, MD: Rowman and Littlefield.

Kelly, K, Odden, A., Milanowski, A., & Heneman, H. (2000). *The motivational effects of school-based performance awards.* CPRE Policy Briefs. Retrieved from http://www.cpre.org/Publications/rb29.pdf

Kennedy, M., Ball, D., & McDiarmid, G. (1993). *A study package for examining and tracking changes in teachers' knowledge* (Technical Series 93-1), East Lansing, MI: NCRTL.

Khattri, N., & Sweet, D. (1996). Assessment reform: Promises and challenges. In M. Kane & R. Mitchell (Eds.), *Implementing performance assessment: Promises, problems, and challenges* (pp.1–21). Hillsdale, NJ: Erlbaum.

Koretz, D., & Barron, S. (1998). *The validity of gains on the Kentucky instructional results information system* (Report No. MR-1014-EDU). Santa Monica, CA: RAND.

Koretz, D., McCaffrey, D., & Hamilton, L. (2001). *Toward a framework for validating gains under high-stakes conditions* (CSE Tech. Rep. No. 551). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Koretz, D., Linn, R. L., Dunbar, S. B., & Shepard, L. A. (1991, April). The effects of high-stakes testing: Preliminary evidence about generalization across tests. In R. L. Linn (chair), *The Effects of High Stakes Testing.* Symposium presented at the annual meeting of the American Educational Research Association and the National Council on Measurement in Education, Chicago.

Ladd, H. F. (2017). No child left behind: A deeply flawed federal policy. *Journal of Policy Analysis and Management, 36*, 461-469. https://doi.org/10.1002/pam.21978

LeCompte, M., & Preissle, J. (1993). *Ethnography and Qualitative Design in Educational Research* (2nd ed.). London: Academic Press.

Leithwood, K., Steinbach, R. & Jantzi, D. (2002). School leadership and teachers' motivation to implement accountability policies. *Education Administration Quarterly, 38*, 94-119. https://doi.org/10.1177/0013161X02381005

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16. https://doi.org/10.3102/0013189X029002004

Linn, R. L. (2004). Accountability models. In S. H. Fuhrman, & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 245–273). New York: Teachers College Press.

Lipman, P. (2004). *High stakes education: Inequality, globalization, and urban school reform.* New York: Routledge. https://doi.org/10.4324/9780203465509

Louis, K. S., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation and Policy Analysis, 27*, 177-204. https://doi.org/10.3102/01623737027002177

Louis, K., Leithwood, K., Wahlstrom, K., & Anderson, S. (2010). *Investigating the links to improved student learning.* Washington, DC: Wallace Foundation.

Lyons, J. E., & Algozzine, B. (2006). Perceptions of the impact of accountability on the role of principals. *Education Policy Analysis Archives, 14*. https://doi.org/10.14507/epaa.v14n16.2006

Malen, B., & Rice, J. (2008). A framework for assessing the impact of educational reforms on school capacity. Insights from studies of high-stakes accountability initiatives. In J. King Rice, J.K., & C. Roellke, (Eds.), *High stakes accountability: Implications for resources and capacity.* Charlotte, NC: Information Age Publishing.

Marshall, C. (1998). Bridging the chasm between policy makers and educators. *Theory into Practice*, *27*, 98–105. https://doi.org/10.1080/00405848809543337

May, P. J. (2012). Policy Design and Implementation in in B.G. Peters & J. Pierre (Eds.) *Handbook of Public Administration* (2nd ed.; pp. 279-291). London: SAGE. https://doi.org/10.4135/9781446200506.n18

McDermott, K. A. (2007). "Expanding the moral community" or "blaming the victim"? The politics of state educational accountability policy. *American Educational research Journal, 44*, 77-111. https://doi.org/10.3102/0002831206299010

McDonnell, L. M. (1994a). *Policymakers' views of student assessment.* Santa Monica, CA: RAND. https://doi.org/10.1037/e652792011-001

McDonnell, L. M. (1994b) Assessment policy as persuasion and regulation. *American Journal of Education, 102,* 394-420. https://doi.org/10.1086/444080

McDonnell, L. M. (2013). Educational accountability and policy feedback. *Educational Policy, 27*, 170-189. https://doi.org/10.1177/0895904812465119

McGill-Franzen, A., & Allington, R. L. (1993). Flunk 'em or get them classified: The contamination of primary grade accountability data. *Educational Researcher, 22,* 19-22. https://doi.org/10.3102/0013189X022001019

McLaughlin, M. W. (1991). Test-based accountability as a reform strategy. *Phi Delta Kappan*, *73*, 248-51.

McLaughlin, M. W. & Talbert, J. E. (1993). *Contexts that matter for teaching and learning.* Stanford, CA: Center for Research on the Context of Secondary School Teaching.

McLaughlin, M. W. & Talbert, J. E. (2001). *Professional communities: and the work of high school teaching.* Chicago: University of Chicago Press.

Metz, M. H. (1990). Real school: A universal drama amid disparate experience. In M. E. Goertz & D. E. Mitchell (Eds.), *Education politics for the new century: The twentieth anniversary yearbook of the Politics of Education Association* (pp. 75-91). London: The Falmer Press.

Meyer, J., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, 83*, 340-63. https://doi.org/10.1086/226550

Meyer, J. W., Scott, W. R., & Deal, T. E. (1981). Institutional and technical sources of organizational structure: Explaining the structure of educational organizations. In H. D. Stein (Ed.), *Organization and the human services: Cross-disciplinary reflections* (pp. 151-179). Philadelphia: Temple University Press.

Mintrop, H. & Trujillo, T. (2007). The practical relevance of accountability systems for school improvement: A descriptive analysis of California schools. *Educational Evaluation and Policy Analysis, 29*(4), 319-352. https://doi.org/10.3102/0162373707309219

Mitchell, K. J. (1996). *Reforming and conforming: NASDC principals discuss school accountability systems.* Santa Monica, CA: RAND.

Moe, T. M. (2003). Politics, control, and the future of school accountability. In P. E. Peterson & M. R. West (Eds.), *No child left behind? The politics and practice of school accountability* (pp. 55–79). Washington, DC: Brookings Institution Press.

Neal, D., & Schanzenbach, W. D. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics, 92*, 263-283. https://doi.org/10.1162/rest.2010.12318

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools.* Cambridge, MA: Harvard Education Press.

Ogawa, R. T. (1992). Institutional theory and examining leadership in schools. *International Journal of Educational Management, 6*, 14-21. https://doi.org/10.1108/09513549210014169

Orton, J. D. & Weick, K. E. (1990). Loosely coupled systems: A reconceptualization. *Academy of Management Review, 15*, 203-223.

O'Toole, L. J. R. Jr. (1995). Rational choice and policy implementation. *American Review of Public Administration, 25*, 43-57. https://doi.org/10.1177/027507409502500103

O'Toole, L. J. R. Jr. (2000). Research on policy implementation: Assessment and prospects. *Journal of Public Administration Research and Theory, 20*, 263-288. https://doi.org/10.1093/oxfordjournals.jpart.a024270

Patton, M. (2015). *Qualitative Research & Evaluation Methods.* (3rd ed.). Atlanta GEO: Sage Publications.

Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). *Perceived effects of state-mandated testing.* Programs on Teaching and Learning: Findings from a National Survey. Boston, MA: National Board of Educational Testing and Public Policy. Lynch School of Education, Boston College.

Pfeffer, J., & Salancik, G. (1978). *The external control of organizations: A resource dependence perspective.* New York: Harper & Row.

Placier, P. L., Hager, D. R., & Hull, A. M. (2005). Appearances can be deceiving. In J. G. Petersen & L. D. Fusarelli (Eds.), *Superintendents and school boards in changing times* (pp. 95-116). Charlotte, NC: Information Age Publishing.

Polikoff, M. S., McEachin, A. J., Wrabel, S. L., & Duque, M. (2014). The waive of the future? School accountability in the waiver era. *Educational Researcher, 43*, 45-54. https://doi.org/10.3102/0013189X13517137

Polikoff, M., McEachin, A., Wrabel, S., & Duque, M. (2015). *Grading the No Child Left Behind waivers.* Washington, DC: American Enterprise Institute.

Powell, W. W., & Di Maggio, P. J. (1991). *The new institutionalism in organizational analysis.* Chicago IL: University of Chicago Press.

Powell, W. W., & DiMaggio, P. J. (Eds.). (2012). *The new institutionalism in organizational analysis.* Chicago, IL: University of Chicago Press.

Robertson, J. E. (2003). Teachers' perceptions of accountability at an international school. *Journal of Research in International Education, 2*, 277-300. https://doi.org/10.1177/1475240903002003002

Romberg, T. A., Zarinnia, E. A., & Williams, S. R. (1989). *The influence of mandated testing on mathematics instruction: Grade 8 teachers' perceptions.* Madison, WI: University of Wisconsin, Center for Educational Research, School of Education, and Office of Educational Research and Improvement of the United States Department of Education.

Rowan, B. (1982). Organizational structure and the institutional environment: The case of public schools. *Administrative Science Quarterly, 27*, 259-279. https://doi.org/10.2307/2392303

Rowan, B. (1995). Institutional analysis of educational organizations: Lines of theory and directions for research. In R. T. Ogawa (Ed.), *Advances in research and theories of school management and educational policy* (Vol. 3, pp. 1-20). Greenwich, CT: JIA.

Rowan, B. (2002). The ecology of school improvement: Notes on the school improvement industry in the United States. *Journal of Educational Change, 3*, 283-314. https://doi.org/10.1023/A:1021277712833

Rowan, B., & Miskel, C. G. (1999). Institution theory and the study of educational organizations. In J. Murphy & K. S. Louis (Eds.), *Handbook of research on educational administration.* (2nd ed., pp. 359-383). San Francisco: Jossey-Bass.

Rusch, E. (2005). Institutional Barriers to Organizational Learning in School Systems: The power of silence. *Educational Administration Quarterly, 41*, 83-120. https://doi.org/10.1177/0013161X04269546

Russell, J. L., & Bray, L. E. (2013). Crafting coherence from complex policy messages: Educators' perceptions of special education and standards-based accountability policies. *Education Policy Analysis Archives, 21*(12). https://doi.org/10.14507/epaa.v21n12.2013

Rustique-Forrester, E. (2005). Accountability and the pressures to exclude: A cautionary tale from England. *Education Policy Analysis Archives, 13*(26). https://doi.org/10.14507/epaa.v13n26.2005

Scott, W. R. (1991). Unpacking institutional arguments. In W.W. Powell & P.J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 164-182). Chicago: University of Chicago Press.

Scott, W. R. (2001). *Institutions and organizations: Foundations for organizational science.* (2nd ed.) Thousand Oaks, Ca: SAGE.

Scott, W. R. (2004). Institutional theory: Contributing to a theoretical research program. In K.G. Smith & M.A. Hitt (Eds.), *Great minds in management: The process of theory development* (pp. 460-484). Oxford, UK: Oxford University Press.

Seashore L. K., & Robinson, V. M. (2012). External mandates and instructional leadership: school leaders as mediating agents. *Journal of Educational Administration, 50*, 629-665. https://doi.org/10.1108/09578231211249853

Shepard, L. A., & Dougherty, K. C. (1991, April). *Effects of high-stakes testing on instruction.* Paper presented at the annual meeting of the American Educational Research Association and the National Council on Measurement in Education, Chicago.

Shohamy, E. (2001). *The power of tests.* England: Longman.

Shohamy, E. (2007). Tests as power tools: Looking back, looking forward. In J. D. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 141-152). Ottawa, ON: University of Ottawa Press.

Shohamy, E. (2017). Critical language testing. In E. Shohamy, L. G. Or, & S. May (Eds*.), Language testing and assessment* (pp. 1-15). New York, NY: Springer International. https://doi.org/10.1007/978-3-319-02261-1_26

Singh-Pillay, A., & Alant, B. (2015). Tracing the policy mediation process in the implementation of a change in the Life Sciences curriculum. *African Journal of Research in Mathematics, Science and Technology Education, 19*, 12-22. https://doi.org/10.1080/10288457.2014.1002297

Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher, 20*(5), 8-11. https://doi.org/10.3102/0013189X020005008

Smith, M. L, Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1989). *The role of testing in elementary schools.* Los Angeles, CA: Center for Research on Educational Standards and Student Tests, Graduate School of Education, UCLA.

Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: the mediating role of teachers' zones of enactment. *Journal of Curriculum Studies, 31*, 143-176. https://doi.org/10.1080/002202799183205

Spillane, J. P. (2002). District policy making and state standards: A cognitive perspective on implementation. In A. Hightower, M. S. Knapp, J. Marsh, & M. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 143-159). New York, NY: Teachers College Press.

Spillane, J. P. (2004). *Standards Deviation.* Cambridge, MA: Harvard University Press.

Spillane, J. P. (2013). The practice of leading and managing teaching in educational organisations. In OECD Center for Educational Research and Innovation (Ed.), *Educational Research and Innovation Leadership for 21st Century Learning* (pp. 59-82.). OECD Publishing.

Spillane, J. P., & Burch, P. (2006). The institutional environment and the technical core in K-12 schools: "Loose coupling" revisited. In H.D. Meyer & B. Rowan (Eds.), *The new institutionalism in education* (pp. 89-100). Albany: State University of New York Press.

Spillane, J. P., Gomez, L. M., & Mesler, L. (2009). Notes on reframing the role of organizations in policy implementation: Resources *for* practice, *in* practice. In G. Sykes, B. Schneider, D. N. Plank (Eds.), *Handbook of education policy research* (pp. 409-426). Stanford, CA: National Academy of Education.

Spillane, J. P., & Hopkins, M. (2013). Organizing for instruction in education systems and school organizations: How the subject matters. *Journal of Curriculum Studies, 45*, 721-747. https://doi.org/10.1080/00220272.2013.810783

Spillane, J. P., Parise, L. M., & Sherer, J. Z. (2011). Organizational routines as coupling mechanisms: Policy, school administration, and the technical core. *American Educational Research Journal, 48*, 586-619. https://doi.org/10.3102/0002831210385102

Stecher, B. M. & Borko, H. (2002). *Combining surveys and case studies to examine standards-based educational reform.* (CSE Technical Report 565). Los Angeles: Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Stecher, B. M., & Chun, T. (2001). *School and classroom practices during two years of education reform in Washington State.* CSE Technical Report 550. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Stryker, R. (1994). Rules, resources, and legitimacy processes: Some implications for social conflict, order, and change. *The American Journal of Sociology, 99*, 847-910. https://doi.org/10.1086/230366

Swanson, C. B. (2005). Organizational coupling, control, and change: The role of higher-order models of control in educational reform. In L.V. Hedges & B. Schneider (Eds.), *The social organization of schooling* (pp. 244-269). New York, NY: Russell Sage Foundation.

Tahmasbi, S., & Yamini, M. (2012). Teachers' interpretations and power in a high-stkaes test: A CLA perspective. *English Linguistics Research, 1*(2), 53-61. https://doi.org/10.5430/elr.v1n2p53

Unger, C., Lane, B., Cutler, E., Lee, S., Whitney, J. Arruda, E., & Silva, M. (2008). *How can state education agencies support district improvement?* New York, NY: The Education Alliance, Brown University.

Warren, C. A. B., & Karner, T. X. (2010). *Discovering qualitative methods: Field research, interviews, and analysis.* New York: Oxford University Press.

Watanabe, M. (2007). Displaced teacher and state priorities in a highstakes accountability context. *Educational Policy, 21*(2), 311-368. https://doi.org/10.1177/0895904805284114

Weick, K.E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*, 1-19. Reprinted in J. V. Baldridge & T. Deal (1983). (Eds.), *The dynamics of*

organizational change in education (pp. 15-37). Berkeley, CA: McCutchon.
https://doi.org/10.2307/2391875

Weick, K. E. (1982). Management of organizational change among loosely coupled systems. In P.
Goodman (Ed.), *Change in organizations: New perspectives in theory, research, and practice* (pp. 375-
408). San Francisco, CA: Jossey-Bass.

Weiss, C. (1990). Policy research: Data, ideas or arguments? In P. Wagner, C. Weiss, C., Wittrock, &
H. Wollmann (Eds.), *Social sciences and modern states* (pp. 307-332). Cambridge: Cambridge
University Press.

About the Authors

Lisa Amdur
Tel Aviv University
lisaamdu@post.tau.ac.il
Lisa Amdur teaches at the Division of Foreign Languages at Tel Aviv University and in the MA
TESOL program of TAU International. She also works at the TAU Center for Language Excellence
that provides individual language support to students and faculty.

Irit Mero-Jaffe
Beit Berl Academic College
iritmero@beitberl.ac.il
Irit Mero-Jaffe is a lecturer at Beit-Berl Academic College. She is also a staff member at the Center
for Educational and Social Research at the Beit Berl Academic College within which she executes
evaluation studies.

---

education policy analysis archives

Please send errata notes to Audrey Amrein-Beardsley at audrey.beardsley@asu.edu

Join EPAA's Facebook community at https://www.facebook.com/EPAAAAPE and **Twitter
feed** @epaa_aape.

arquivos analíticos de políticas educativas
conselho editorial