Three Decades of School Failure in Swedish Compulsory School

Cecilia Arensmeier

Department of Humanities, Education and Social sciences, Örebro University, Örebro, Sweden

ABSTRACT
In Sweden, compulsory school grades determine admission to upper-secondary school. The article maps grading outcomes in compulsory school for 1990–2017, when three different grading scales were used, in terms of students' distribution across grading steps. Statistics of grades for all Swedish grade 9 students (all school subjects) are used. Contrary to policymakers' expectations, the results show that a large proportion of students failed to pass compulsory school immediately after a criterion-referenced system with a sharp pass/fail distinction was introduced in the late 1990s. The failure rate has since then remained strikingly constant. Swedish as a second language differs from the main pattern, with a substantially higher failure rate that is increasing over time. This outcome is discussed with reference to grading policy as a matter of social choice.

Introduction

Grading is an essential feature of education, not least in contexts where teacher-assigned grades determine students' prospects for further education (C. Lundahl et al., 2017). Formal institutional aspects of assessment are particularly prominent when politicians decide over educational goals and grading systems (Salomonsen & Andersen, 2014). Sweden is a clear example of both; grades are a "high stakes" issue, and learning objectives and grading systems are politically decided. Over the years, Sweden has also been characterised by recurring political debates about grading (C. Lundahl et al., 2015, p. 772). Three major reforms have been conducted since the mid twentieth century. The comprehensive nine-year compulsory school that was launched in the 1960s used a norm-referenced ranking system without failing grades. This was replaced by a criterion-referenced system with a sharp pass/fail distinction in the 1990s, which, in turn, was revised in the 2010s. Since the 1990s, the admission requirements for the vocational and academic tracks in upper-secondary school have also been linked to passing grades. The requirements were sharpened in the 2010s. These grading reforms were framed very differently, which will be further illustrated below (Arensmeier, in review).

The transfer to criterion-referenced grading brought about a profound change of logic and language (cf. Wikström, 2006). The way the least successful children are spoken of, in terms of grades, changed from students "having bad/low grades" to "being failing or failed students". This illustrates how policy changes can affect who students are or become (Ball, 2015), and brings attention to ethical concerns like "powerful social consequences of assessment policy" (Elwood, 2013, p. 206).

The aim of this article is to map the grading outcome over time in Swedish compulsory school, in terms of how students are distributed across different grading scales, and in relation to this discuss...
Swedish grading policy as a matter of social choice. The empirical material consists of grading statistics for all Swedish grade 9 students (and all school subjects) from 1990 to 2017, during which time three different grading scales were used. The questions asked are:

- Which grades (performances) in the norm-referenced system “transferred” into failing grades with the introduction of a criterion-referenced system in the late 1990s?
- How has the “failure rate” developed over time and are there differences between subjects?
- How can the outcome be discussed if grading policy is regarded as a social and ethical choice?

An important point of departure for the study is an expectation repeatedly raised in the process leading up to the introduction of the first criterion-referenced grading system in the 1990s. Even though a sharp pass/fail distinction was proposed, “virtually all students” were expected to reach passing level in all subjects (Prop., 1992/93:220, 1993, p. 75; SOU, 1992:86, 1992, p. 67). Failure was therefore more or less considered a non-issue (Arensmeier, in review). As will be shown in more detail in this article, this expectation was proven wrong. The results “diverged from the intentions of the reformers” (Cortell & Peterson, 2001, p. 772). Unintended outcomes of this kind are not in themselves necessarily good or bad in normative terms. Attention to ethical aspects might, however, encourage discussion about desired or unintended social consequences of policy choices.

The Swedish Context

Compulsory School

Swedish compulsory school is equivalent to primary and lower-secondary education. It consists of three stages, years 1–3 (lower stage), years 4–6 (middle stage) and years 7–9 (upper stage). Grades are currently given in the final year of the middle stage and every semester of the upper stage. Regular meetings with students, parents and teachers ("development talks") are held throughout compulsory school. Mandatory national tests are conducted in year 3 (Math, Swedish), year 6 (English, Math, Swedish) and year 9 (as in year 6, plus one science subject and one social studies subject). Having to repeat a grade is very rare, even among students with numerous failing grades.

The introduction of comprehensive, nine-year compulsory schooling in the 1960s turned the Swedish school system into one of the most progressive in the world, with a kind of role-model status (Oftedal Telhaug et al., 2006). Reforms in the late 1980s and 1990s meant substantial changes. Decentralisation, free school choice, and the right for different organisations – including for-profit businesses – to run schools in a publicly financed voucher system, were implemented. This transformed Swedish basic education into one of the most market-oriented systems in the world (Imsen et al., 2017; L. Lundahl et al., 2013; Oftedal Telhaug et al., 2006).

Today, the main responsibility for running schools lies with the municipalities, along with a number of different school providers. Equality and equity remain core values for basic education, which not least is visible in the national school legislation and curricula, which have continuously directed the schools. Over the last decade, the state has also regained some power and control through legislation, regulation and inspection (Bergh et al., 2018; Lilja, 2014; C. Lundahl & Waldow, 2009; Rönnberg, 2014). This has, however, taken place within a framework of management by objectives (Imsen et al., 2017), an ideal very visible in the grading system.

1The same grading scale, including a sharp pass-fail distinction, is used in upper-secondary school. Here, ‘school failure’ has considerable consequences for chances in the labour market (Helgøy et al., 2019). Upper-secondary school, however, is not included in this study.
Grading System Reforms

The grading system has always been controlled by the state. The three reforms conducted since the mid twentieth century, were framed very differently though (Arensmeier, in review). The norm-referenced grading system, first implemented in primary school and then chosen for the comprehensive compulsory school in the 1960s, was embedded in a desire to make society more egalitarian. The main goal was to enhance uniformity in grading, so that grades could be used to fairly rank applicants in the competition for the next educational level, and admissions tests would be unnecessary. This was viewed as particularly beneficial for gifted students from poor backgrounds. Norm-referenced grading was used until 1997, based on a 5-grade numerical scale, with 5 as the best grade. The system rested on an assumption of normal distribution of performances within the national population of each school year, and was supported by national standardised tests in some subjects. Grade 3 was defined as the normal/average performance. The lowest grades intentionally had no other official meaning than that they indicated a low ranking compared to the national population (Arensmeier, in review).

In the 1990s, a criterion-referenced grading system was introduced, during a reform period characterised by decentralisation, management by objectives and marketisation (Andersson, 2005; Tholin, 2006; Wikström, 2006; Arensmeier, in review). Grades were intended to reflect the degree to which learning objectives had been achieved. By relating grades to learning objectives and by introducing a pass/fail distinction and terminology, with reference to minimum passing levels, it was argued, grades could serve as a results measure for school quality. Moreover, the system was also thought to have pedagogical merits that would enhance learning, which was seen as especially beneficial for the weakest students (Prop., 1992/93:220, 1993; SOU, 1992:86, 1992, p. 86). To emphasise school and teacher accountability for learning outcomes, no failing grades were awarded in compulsory school; students who did not reach the passing level received no grade at all in that subject. The passing grades used word-labels, literally meaning “approved” (godkänd, G), “well approved” (väl godkänd, VG), and “very well approved” (mycket väl godkänd, MVG), but often translated as “pass”, “pass with distinction” or “pass with special distinction”.

As already mentioned, it was clearly assumed that virtually all students would pass in all subjects. At the same time, or perhaps as a result of this assumption, the passing levels as such were not discussed in any thorough way during the reform preparations, and the empirical testing was also very poorly conducted (Arensmeier, in review). However, a quite large proportion of students failed in one or more subjects immediately after the introduction of the new grading system (see Figure 5), but this did not cause any debate about the system as such. Instead, attention was directed to shortcomings in schools and teaching (Mickwitz, 2015; Skolverket, 2001).

A decade into the twenty-first century, the grading system was reformed again, in the context of a political debate about a Swedish school crisis, visible in the large proportions of failing students, and – in particular – decreasing results in international studies like the Programme for International Student Assessment (PISA). The criterion-referenced system was retained, but a scale with more steps and an explicit failing grade was gradually introduced from 2011, along with grades from school year 6 instead of year 8. Earlier assessment and more national testing were argued to be a remedy for the crisis, since they could help identify students in need of more support (Prop., 2009/
Table 1. Characteristics of grading scales and reforms in Swedish compulsory school.

| Grading system | Grading scale (low to high) | Selection for upper-secondary school | Framing of reform |
|----------------|-----------------------------|--------------------------------------|-------------------|
| Norm-referenced 1960s to 1997 | Ranking on a numerical scale 1–5 (--; if low attendance) | Grade point average for competition ↢ compensation possible (good grades can counterbalance low grades in some subjects) | Fair grading can help fight inequality No one should fail compulsory school; lowest grade a relative measure |
| Criterion-referenced Mid 1990s to 2012 | No grade if learning objectives are not reached (0 qp) Three passing grades: - Approved (Godkänd) (10 qp) - Well approved (Väl godkänd) (15 qp) - Very well approved (Mycket väl godkänd) (20 qp) | Total of qualification points for competition Pass in the three ‘core subjects’ Swedish, Math, English always required for admission ↢ no grade in these subjects closes the door to regular secondary education | Grades a result measure for school accountability Virtually every student expected to pass + criterion-referenced grading benefits weakest students |
| Criterion-referenced 2013- | (--; if low attendance/not sufficient information to assess (0 qp) One failing grade: F (0 qp) Five passing grades: - E (10 qp) - D (12.5 qp) - C (15 qp) - B (17.5 qp) - A (20 qp) | Total of qualification points for competition Pass in the three core subjects + additional subjects required for admission - 3+5 for vocational track - 3+9 (of relevance to programme) for academic track ↢ F-grades can close door to regular secondary education | More and earlier grading can combat the school crisis Confirming pass levels At least ‘making an effort’ acknowledged by explicit failing grade (F); failing normalised |

Source: Based on Arensmeier (in review)

The admission requirements for upper-secondary school were sharpened at the same time (see Table 1 below). The passing levels as such were not questioned, but instead were confirmed. The changes made in national curricula focused on establishing clearer learning objectives and assessment criteria, not on changing the level of difficulty. Introducing an explicit failing grade in compulsory school was motivated by a need to differentiate between students who were at least trying, and students who were absent from school and could not be graded. Due to the ongoing PISA crisis, any questioning of the appropriateness of existing pass levels was not politically viable. The explicit failing grade was instead normalised as a grade among others (Arensmeier, in review).

### Admission to Upper-Secondary School

Under all three grading scales, compulsory school grades have determined admission to upper-secondary school, which in Sweden includes a vocational and an academic (university preparatory) track, each divided into different study programmes. In the norm-referenced grading system, grade point average served as selection tool in cases of competition. In the criterion-referenced systems, grades are converted into qualification points (qp), which are used for ranking. Vital to this estimate is that failing grades give 0 qp, while the lowest passing grade equals 10 qp. Additional points are then given for every grading step, up to 20 qp for the highest grade. The (re)introduction

---

4Reforms in the 1970s brought together vocational and academic tracks in the same upper-secondary school, gymnasieskolan, which also expanded to be a viable educational path for all youth.
of a fail-pass distinction also came, however, with minimum requirements for admission to upper-secondary school, based on passing grades in certain subjects. The latest changes have toughened these requirements. Table 1 summarises the main characteristics of the three grading scales, and how policy changes have been framed.

**Assessment Policy as a Matter of Social Choice**

As shown, political decisions determine which grading systems and scales are used in Swedish schools. Political choices depend on many factors, not least how problems are thought about and framed (Bacchi, 2009). Grading is a form of assessment. Messick (1981) brings attention to ethical aspects of assessment and discusses tests as a matter of social choice, potentially giving rise to both intended and unintended outcomes and tensions. He therefore finds it important to examine the use and interpretation of tests, by looking at their social context, purpose and consequences. For example, Messick underlines that tests intended to measure “mastery” and/or capture “minimum competencies”, include two models of interpretation: a cumulative scale, where a higher score indicates higher competence; and a categorical scale, which classifies scores below a certain level as non-mastery or failure. Additionally, since agreement about evaluative terms such as “mastery” or “basic skills” is rare, interpretation becomes even more disputable (Messick, 1981, p. 25ff.).

On a general level, tests also create winners and losers, success and failure, which can “have far-reaching implications for individuals and educational systems” (Shohamy, 2001, p. 374). In an analysis of how tests are interpreted and used for academic selection to secondary schools in Northern Ireland, Elwood (2013) underlines the importance of analysing assessment systems thoroughly before they are implemented, since they can have “severe social and educational consequences for the children who experience them” (p. 216).

These examples relate to high-stakes test. However, the same kind of ethical considerations about policy choices can be raised in contexts where grading is both politically regulated and a high-stakes issue, such as in determining admission to higher educational levels. The concern that Messick, Shohamy and Elwood direct to consequences for children is therefore relevant to the Swedish case. All the tensions that Messick highlights are visible in the current grading system. It rates performances cumulatively, includes a categorical pass/fail distinction, and is based on outcome labels like “approved”, all of which are specified in the politically decided curricula. Unlike tests, school leaving-certificate grades do not depend on one single performance, but are based on several assessments over a longer period of time. On the one hand, this might make the ethical concerns less pressing. But on the other hand, the use of failing grades might raise even greater concerns. Grades of this kind, with their wider scope, are more definite and cannot be “rationalised” with reference to temporary factors like poor preparation or having a bad day.

In a historical review, Lysne (2006) brings attention to how earlier negative grades in Denmark and Norway could have devastating effects on students’ grade point averages. This was particularly troublesome for slow learners, but even stronger students could have their records ruined by one or two very low marks. A quasi-experimental study from Sweden (data from norm-referenced system), where one group of students was graded in both grades 6 and 7, and the other only in grade 7, also points to negative effects for some students. In contrast to other students, where no effect on later achievements was noted, students with low cognitive ability received lower grades in grade 7 if they also had been graded the previous year (Klapp et al., 2016). The same goes for the longer term. For students with low cognitive ability, being graded in grade 6 also has negative consequences for the grade level in grades 8–9 and for the possibility to complete upper-secondary school (Klapp, 2015). In a system like the current Swedish one, where a single failing grade might prevent admission to upper-secondary school, the implications can be even more devastating for students, both in terms of self-image (Löfgren et al., 2019; Räty et al., 2004) and their chances in the labour market (Helgøy et al., 2019).
**Method and Material**

To map how different grading scales have “sorted” students over time, the article uses grade statistics from Statistics Sweden. The material covers all subjects for the entire population of students in grade 9 in Swedish compulsory schools over the years 1990–2017. This amounts to between 95,000 and 125,000 students every year.  

The development is illustrated by descriptive area charts (made in Excel), which display the distribution of students across the grading steps used at different times. The distributions under the three different scales are presented next to each other, to provide a graphic visualisation of outcomes over time. As Bryman (2012) argues, graphs of this kind of extensive data are usually quite easy to understand and interpret. The descriptive approach is also in line with Proches (2016) plea for the merits of description in research. The descriptive outcome, however, is also related to the policy development, and linked to a thorough qualitative analysis of policy documents (Arensmeier, in review).

As shown, the logics of the systems differ, and of course it is not adequate to make straightforward comparisons of the outcomes. However, it is possible to relate the patterns occurring during different periods to each other. What is compared is the distribution of students across each of the grading scales used at different times. The scales also share the essential feature that they grade performances from lower to higher, with reference to a norm group or to learning objectives and assessment criteria (cf. SOU, 2016:25, 2016, appendix 3, p. 149). For this purpose, dissimilarities in the principles used for ranking, number of grading steps, and grade labels, are of less importance. Another argument for relating the outcomes under different systems to each other is that the all grade 9 students in Sweden during the current years are included. This ensures that the entire range of school performances is always included. Further, it is reasonable to believe that neither the general performances of grade 9 students as a group, nor teachers’ assessment practices, change very much from one year to the next, even when the grading scale does. Needless to say, caution in interpretation is still called for.

Compilations of the statistics have been made for the seventeen subjects that have been part of compulsory school during the entire period of interest. The article summarises the main patterns, and illustrates the outcomes in more depth for five different subjects: history, music, physics, Swedish and Swedish as a second language (SweSL). These represent the whole range of subjects from social studies and science to arts and languages.

The descriptive statistics are then discussed with reference to the policy processes framing the reforms, as they appear in reports from government commissions of inquiry (presented in Statsens Offentliga Utredningar (SOU) – Official Reports of the Swedish Government), reports of the Ministries, (Departementsserien (Ds)), government bills, parliamentary committee reports and minutes from parliamentary debates. The policy process and documents have been analysed in depth elsewhere (Arensmeier, 2019; Arensmeier, in review).

**Results: Grading Outcomes 1990–2017**

The section starts with an overview of the grading outcomes in 17 subjects, focusing on the years immediately following grading reforms (Table 2). This is followed by a closer look at history, music,
Table 2. Failure rates after introduction of criterion-referenced grading.

| Subjects                      | 1998–2000 not attained pass | Proportion not passed compared to norm-referenced scale 1–5 (1997) | 2013–15** not attained pass | Proportion ‘not passed’ compared to previous scale (2010–12) |
|-------------------------------|------------------------------|------------------------------------------------------------------|-----------------------------|-------------------------------------------------------------|
| Languages                     | Swedish  4–5%                | All grade 1 + about 5% of grade 2*                               | Swedish 4%                  | Similar, slight decrease in modern languages                |
|                               | English 6–7%                 | All grade 1 + about 15% of grade 2                               | English 7–8%                |                                                             |
|                               | Modern  4–6%                 | All grade 1 + about 15% of grade 2                               | Modern 4–5%                 |                                                             |
| Aesthetic + Practical subjects| About 4–8% (highest in PE)   | All grade 1 + about 15% of grade 2 (a bit higher in PE)          | About 5–9% (highest in PE)  | Similar or slight increase                                  |
| Social studies subjects       | About 7–9%                   | All grade 1 + 10–20% of grade 2                                  | About 8–9%                  | Similar                                                     |
| Math                          | Math 6–8%                    | All grade 1 + about 10% of grade 2*                              | Math: 9–11%                 | Similar or slight increase                                  |
| Science                       | Technology 6–7%              | All grade 1 + about 15–25% of grade 2                            | Technology 7–8%             |                                                             |
|                               | Biology 8–10%                | All grade 1 + 10–20% of grade 2                                  | Biology 9–10%               |                                                             |
|                               | Chemistry, Physics 9–11%     | All grade 1 + 10–20% of grade 2                                  | Chemistry, Physics 11–12%   | Similar                                                     |
| Swedish as Second Language    | About 21–25% (increasing)    | All grade 1 + 70–90% of grade 2                                  | About 31–37% (increasing)   | Leap upwards, continuous increase                          |

* Until 1997, English and Math had a basic and an advanced track, and few students took the subjects without tracking. The reference group from 1997 is therefore limited. Tracking was abandoned in 1998.
** The proportion of students not attaining the goals (F-grades, dashes, adapted studies) in 2016–17 were notably affected by the refugee situation of 2015, which added quite large numbers of refugee teenagers to Swedish classrooms.

physics, Swedish and Swedish as a second language, where Figures 1–4 display the distributions across all grade steps in these subjects. Lastly, the proportion of students without grade(s) or who have not reached the goals in one or more subjects between 1990 and 2019 is displayed (Figure 5).
Mapping students’ distribution across grading steps reveals both a general pattern and some differences between subjects. Table 2 summarises the proportions of failing /low grades during the use of different scales and compared to previous scales. The sharp pass/fail distinction, that came with the first criterion-referenced grading system in the late 1990s, resulted in a proportion of “failure” that amounted to approximately 4–6 per cent in languages – except SwSL – up to about 10 per cent in some science subjects. In SwSL, the proportion was over 20 per cent. The shares were approximately equal to the shares of students who, in the previous norm-referenced grading system, received grade 1 and, additionally, a small to somewhat substantial proportion of students who received grade 2. The proportion of failing students often increased a bit in the first years, but then stabilised (Figures 1–4).

As indicated by the similarity in proportion of failing students in Table 2 and Figures 1–4, and in line with the intentions, the pass/fail boundary did not change when the new criterion-referenced scale (A–F) was introduced in the 2010s. The proportion of failing students remained almost exactly the same. Further, there are no signs that changes in the grading scale (and other measures taken) have contributed to reducing the proportion of failing students over time.

As intended in that system, there is a notable stability in students’ distribution across the grading steps during the last years of norm-referenced grading. The exception is the final year, 1997, where a slight increase in the highest grade is seen in many subjects. In contrast, signs of grade inflation – that is, a trend where teacher-assigned grades rise in an unjustified manner – are visible in many subjects under the two criterion-referenced systems, most markedly under the most recent scale
That the pattern can be interpreted as grade inflation has been shown in other studies (Björklund et al., 2010; Gustafsson et al., 2014; Tyrefors Hinnerich & Vlachos, 2016; Vlachos, 2018). One notable exception from the main pattern is Swedish as a second language (SwSL). From the start, the proportion of “failure” was much higher in SwSL than in all the other subjects. The proportion was equivalent to almost all students who in 1997 received grade 1 or grade 2. It has also continuously increased during the use of both criterion-referenced grading scales.

The proportion of students who take SwSL instead of regular Swedish increases during the studied period, from about 3–4 per cent in the early 1990s to more than 15 per cent in the last few years. The development mirrors immigration.
The proportions of failing grades after system changes (columns 2 and 4 in Table 2), compared to the pattern for the lowest grades in the previous system (columns 3 and 5), show many similarities and some differences between subjects. Additionally (but not shown in the Table), under all three grading scales, the highest grade levels are found in arts/practical subjects, modern languages (optional), English (advanced track 1990–97), and Swedish (under the criterion-referenced systems). The lowest grades are visible in math (basic level 1990–97), chemistry and physics, and – at a level of its own under the criterion-referenced scales – Swedish as a second language.

A closer look at the grade distribution in five subjects adds some nuance to the picture. The outcome for history illustrates the stability of failing grades, after a slight increase in the first years after the introduction in 1998. The graphs also display signs of grade inflation under the two criterion-referenced systems. It also appears as if the highest grade was most difficult to receive in the norm-referenced system. The arrival of many new refugees, with little or no general schooling, and no experience of Swedish school, is also particularly visible from 2016. The same goes for most other subjects. The national grade point average for history was a bit above the 3.0 mean during the last years of norm-referenced grading. With values around 12.4–12.9 it is approximately as far from the theoretical max–min values in the criterion-referenced systems.

The overall pattern for music is similar to that for history. However, the failure rate under the criterion-referenced scales is a bit lower, among the lowest of all subjects. The signs of grade inflation are also apparent under both criterion-referenced scales. The national grade point average for 1990–97 was well above 3.0. With reference to respective max–min values, the qualification point average under the first criterion-referenced scale started out somewhat lower, but has since gradually increased to levels that are comparatively higher (in terms of closeness to the max-value). In recent years, it is approaching 14.0.

Compared to most other subjects, the pattern for physics reveals a higher failure rate when criterion-referenced grading was introduced in the late 1990s. After a slight increase in the first few years, the level stabilised here as well. Grade inflation in physics is only indicated after the introduction of the A–F scale. The national grade point average in the 1990s was somewhat above 3.0, slightly lower than for history. Under the criterion-referenced scales, average qualification points have become somewhat lower, comparatively speaking, with values under or around 12.0.

The patterns in Swedish and Swedish as a second language (SwSL) are completely different. The results for Swedish resemble those for music, with low failure rates, clear signs of grade inflation, and increasing average qualification points. No dip occurred in 2016, which is logical since newly arrived refugees take SwSL. The grade point average in Swedish for 1990–97 was similar to that for history. In 1998, the average qualification points were at a roughly similar level (12.7), but have since gradually increased to exceed 14.0 in 2017.

The outcome for SwSL is very different. Compared to all other subjects, a much larger proportion of students immediately failed SwSL after the introduction of criterion-referenced grading. An increase in low grades can be noted already during the last years of the norm-referenced system, but the effect of the introduction of criterion-referenced grading differed drastically from that for other subjects, with a much higher failure rate. In the first years, the proportion was almost equivalent to all grades below 3 in 1997. This indicates that, compared to other subjects, SwSL became more “demanding” than it was before. Contrary to other subjects, the failure rate for SwSL has also continuously increased. This is likely related to migration and an increased number of students taking the subject. The same trend is also visible in the national grade point averages. In the first
years of the 1990s, the average was in line with that for Swedish, but it decreased to just under 3.0 in 1997. A clear leap downwards, in relation to max–min values, occurred in 1998 with an average qualification points score of about 9.4. A constant decrease has since followed, to 8.5 in 2013, and well under 7 in 2016–17.

Figure 5 shows the development of students without grades in one subject or more. Until 1998, this meant no grade at all in at least one subject, and from 1998 it means students who have failed to meet the passing level in one subject or more.

Since the introduction of criterion-referenced grading, one out of every four or five students has thus consistently “failed” Swedish compulsory school. Furthermore, failing grades in too many or the “wrong” subjects can prevent a student from continuing directly to regular upper-secondary school (see Table 1).

Discussion: Persistence and Grading Policy as a Social Choice

Not meeting the minimum requirement for passing has implications for both individual students and the educational system (Shohamy, 2001). One single failing grade places 15–16 year olds in the category of students who leave compulsory school “without complete grades”. With some variation between years, the proportion of students not meeting the admissions requirements for regular upper-secondary school was about 8–12 per cent during 1998–2012. Since the sharpening of admission standards in 2013, it has varied between 12 and 17 per cent (Skolverket, 2019a).13

As shown, the proportion of students who do not pass is strikingly stable in all subjects from the turn of the century and onwards. This persistence is particularly notable in the light of the grade inflation that is indicated during the same time-period (Björklund et al., 2010; Gustafsson et al., 2014; Tyrefors Hinnerich & Vlachos, 2016; Vlachos, 2018). The only deviating subject is SwSL, where the failure rate has constantly increased in parallel with more and more students taking the subject. Expressed somewhat drastically, from one year to the next, the grading reform of the 1990s transformed low-performing students into failing students, and this was especially the case for students with an immigrant background. The logic of failure was also enhanced by the fact that one single weak performance (grade) was enough to categorise a student’s performance in compulsory school as unsatisfactory.

The high hopes for and confidence in the “social choice” (Messick, 1981) of introducing a criterion-referenced grading system with a sharp pass/fail distinction in the 1990s effectively prevented political debate about the risk of constructing failing students (Arensmeier, in review). More notable is that ethical considerations of this kind were also lacking when the grading scale was again reformed in the 2010s, despite the fact that by then the expectations of the 1990s had been proven wrong. Virtually all students did not meet the passing level in all subjects. On the contrary, as shown in this study, from the very beginning, quite a few students failed compulsory school in a strikingly stable pattern. The reform of the 2010s was launched at a time when the political discourse about education was permeated by a “knowledge crisis” in Swedish schools. This probably explains why neither the categorical thinking and terminology, nor the passing levels as such, were given any attention when the grading system was again reformed. The passing levels were instead reaffirmed (Arensmeier, in review). The government bill provides a clear example of this:

Grade E will be given for knowledge that meets the lowest acceptable knowledge requirements. […] The proposed scale is thus directly transferable to the grading step that in the current system denotes approved results. (Prop., 2008/09:66, 2009, p. 14f.)

---

13 There are currently 18 national programmes, 12 of which are vocational and 6 higher-education preparatory. Students who do not meet the admissions requirements are offered to attend an introductory programme, which is given in five different forms. The aim of these programmes is to give students who are ineligible for national programmes an education that is individually adapted, often with the ambition that students can eventually qualify for regular programmes.
Arguments that F grades should at least be assigned some qualification points were also dismissed, which also emphasizes the categorical logic of pass and fail, approved and non-approved performances:

It is essential that the value scale of the grading system signals that it is important to achieve approved results. (Prop., 2008/09:66, 2009, p. 16)

A strong discourse of accountability has dominated the Swedish political discussion about criterion-referenced grading, and it has been shown that this has contributed to a questioning of teachers’ professionalism (Mickwitz, 2015). The chosen grading system and its rationale also, of course, have implications for children (Elwood, 2013). They are the ones who are failing (Klapp, 2015; Klapp et al., 2016; Löfgren et al., 2019), and it is their prospects for further education and success in the labour market (Helgøy et al., 2019) that are at stake. Paths to upper-secondary school still exist for students who lack eligibility after compulsory school, but having to take these special roads due to "failure" is still often experienced as degrading.

**Concluding Remarks**

That the passing levels introduced in the 1990s seem to have been set significantly higher than foreseen has so far hardly been considered in the Swedish policy debate, which seems more concerned with matters like how students can be better supported (Dir., 2017:88), equivalency problems (Skolverket, 2019b), and unclear assessment criteria (e.g., Måhl, 2014). Recently, some media and research attention can be noted however. For example, a study of school failure in grade 6 indicates that some students probably lack some of the cognitive abilities needed to pass, and this cannot be remedied even with substantial long-term pedagogical support (Lindblad et al., 2018). In other words, given the current knowledge requirements and the sharp passing level, some children appear to be doomed to fail compulsory school, no matter how much they learn and develop within the scope of their capacity.

Messick suggests that alternative proposals or counterproposals, with regard to the interpretation and use of assessment tools, can direct attention to vulnerabilities and fuel open debate (Messick, 1981). In line with this, it can be asked whether Sweden ought to abandon a grading scale that continuously defines a large proportion of compulsory school students as “failed” or “non-approved”, something that can prevent admittance to both vocational and university-preparatory upper secondary school. These outcomes were neither anticipated nor intended. Nevertheless, both the short – and long-term effects of these policy choices can be devastating for the children concerned. The contrast to the Swedish self-image of a having a humane and egalitarian school system is striking. That schools and teachers are constantly portrayed as inadequate is another implication.

Returning to the norm-referenced system, changing the passing levels, omitting the fail label for grade F, abandoning high-stakes grading in compulsory school altogether, or changing the requirements for admission to upper-secondary school are all potential policy choices that could be made. Retaining the current system, grading scale and rationale is also possible of course. However, as indicated by the results presented in this article, this would probably mean that a quite large proportion of students will continue to fail compulsory school, or, put differently, that schools and teachers will continue to fail in their mission.

**Acknowledgements**

The research was supported by the Swedish Research Council (Vetenskapsrådet) under Grant 721-2013-2212.

**Disclosure Statement**

No potential conflict of interest was reported by the author(s).
**Funding**

The research was supported by the Swedish Research Council under [grant number 721-2013-2212].

**ORCID**

Cecilia Arensmeier  [http://orcid.org/0000-0003-3408-6829](http://orcid.org/0000-0003-3408-6829)

**References**

Andersson, H. (2005). Holding the axe: Equivalence, selection and justice in the reform of school-based assessment in Sweden. *Pedagogy, Culture & Society, 13*(2), 143–158. [https://doi.org/10.1080/09500780500200221](https://doi.org/10.1080/09500780500200221)

Arensmeier, C. (2019). *Grundskolebetyg under tre betygsystem: Fördelning på betygsspegel och genomsnittliga ämnesbetyg 1990-2017*. [Grades in compulsory school under three grading systems: Distribution across grades and average subject grades 1990-2017]. Örebro university.

Arensmeier, C. (in review). Institutionalizing school failure: From abandoning to reintroducing a failing grade – the rationales behind three Swedish grading reforms.

Bacchi, C. L. (2009). *Analysing policy: What’s the problem represented to be?* Pearson.

Ball, S. J. (2015). What is policy? 21 years later: Reflections on the possibilities of policy research. *Discourse: Studies in the Cultural Politics of Education, 36*(3), 306–313. [https://doi.org/10.1080/01596306.2015.1015279](https://doi.org/10.1080/01596306.2015.1015279)

Bergh, A., Löfdahl Hultman, A., & Englund, T. (2018). Local enactment of the Swedish ‘advanced teacher reform’. *Journal of Curriculum Studies, 1-16*. [https://doi.org/10.1080/00220227.2018.1436195](https://doi.org/10.1080/00220227.2018.1436195)

Björklund, A., Fredriksson, P., Gustafsson, J.-E., & Öckert, B. (2010). *Den svenska utbildningspolitikens arbetsmarknadseffekter: Vad säger forskningen?* [Labour market effects of Swedish education policy: what does the research say?]. Institutet för arbetsmarknadspolitisk utvärdering.

Bryman, A. (2012). *Social research methods*. Oxford University Press.

Cortell, A. P., & Peterson, S. (2001). Limiting the unintended consequences of institutional change. *Comparative Political Studies, 34*(7), 768–799. [https://doi.org/10.1177/01044001034007003](https://doi.org/10.1177/01044001034007003)

Dir. [Government directive to a commission of inquiry]. (2017:88). [Bättre möjligheter för elever i de obligatoriska skolformerna att nå de kunskapskrav som minst ska nås. [Better possibilities for students in mandatory school forms to achieve the minimum knowledge requirements]]. The Swedish Government.

Elwood, J. (2013). Educational assessment policy and practice: A matter of ethics. *Assessment in Education: Principles, Policy & Practice, 20*(2), 205–220. [https://doi.org/10.1080/0969594X.2013.765384](https://doi.org/10.1080/0969594X.2013.765384)

Gustafsson, J.-E., Cliffordson, C., & Erickson, G. (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan: Problem och möjligheter*. [Equivalency in knowledge assessment in and about Swedish school: problems and possibilities]. SNS förlag.

Helgøy, I., Homme, A., Lundahl, L., & Rönning, L. (2019). Combating low completion rates in Nordic welfare states: Policy design in Norway and Sweden. *Comparative Education, 1–18*. [https://doi.org/10.1080/03050068.2019.1619328](https://doi.org/10.1080/03050068.2019.1619328)

Imsen, G., Blossing, U., & Moos, L. (2017). Reshaping the Nordic education model in an era of e-education. *Comparative Education, 61*(5), 568–583. [https://doi.org/10.1080/03050068.2016.1172502](https://doi.org/10.1080/03050068.2016.1172502)

Klapp, A. (2015). Does grading affect educational attainment? A longitudinal study. *Assessment in Education: Principles, Policy & Practice, 22*(3), 302–323. [https://doi.org/10.1080/0969594X.2014.988121](https://doi.org/10.1080/0969594X.2014.988121)

Klapp, A., Cliffordson, C., & Gustafsson, J.-E. (2016). The effect of being graded on later achievement: Evidence from 13-year olds in Swedish compulsory school. *Educational Psychology, 36*(10), 1771–1789. [https://doi.org/10.1080/01443410.2014.933176](https://doi.org/10.1080/01443410.2014.933176)

Lilja, P. (2014). A quest for legitimacy: On the professionalization policies of Sweden’s teachers’ Unions. *Journal of Education Policy, 29*(1), 86–104. [https://doi.org/10.1080/02680939.2013.790080](https://doi.org/10.1080/02680939.2013.790080)

Lindblad, I., Westerlund, J., Gillberg, C., & Fernell, E. (2018). Har alla barn i grundskolan förutsättningar att klara nya lärplanens krav? [Can all children in compulsory school achieve the new curriculum goals?] *Läkartidningen, 115*(6), 1–4.

Löfgren, R., Löfgren, H., & Lindberg, V. (2019). *Pupils’ perceptions of grades: A narrative analysis of stories about getting graded for the first time*. *Assessment in Education: Principles, Policy & Practice, 26*(3), 259–277. [https://doi.org/10.1080/0969594X.2019.1593104](https://doi.org/10.1080/0969594X.2019.1593104)

Lundahl, C., Hultén, M., Klapp, A., & Mickwitz, L. (2015). *Betygens geografi: forskning om betyg och summativa bedömningar i Sverige och internationellt*. [The geography of grades: research about grades and summative assessment in Sweden and internationally]. The Swedish research council.

Lundahl, C., Hultén, M., & Tveit, S. (2017). The power of teacher-assigned grades in outcome-based education. *Nordic Journal of Studies in Educational Policy, 3*(1), 56–66. [https://doi.org/10.1080/20020317.2017.1317229](https://doi.org/10.1080/20020317.2017.1317229)
Lundahl, C., & Waldow, F. (2009). Standardisation and ‘quick languages’: The shape-shifting of standardised measurement of pupil achievement in Sweden and Germany. *Comparative Education*, 45(3), 365–385. https://doi.org/10.1080/03050060903184940

Lundahl, L., Arreman, I. E., Holm, A.-S., & Lundström, U. (2013). Educational marketization the Swedish way. *Education Inquiry*, 4(3), 497–517. https://doi.org/10.3402/edui.v4i3.22620

Lyse, A. (2006). Assessment theory and practice of students’ outcomes in the Nordic Countries. *Scandinavian Journal of Educational Research*, 50(3), 327–359. https://doi.org/10.1080/00313830600743565

Mähl, P. (2014, April 23). Ej godkänt! Tummen ned för nya betygen. [Not approved! Thumbs down for the new grades], *Pedagogiska magasinet*.

Messick, S. (1981). *Evidence and ethics in the evaluation of tests*. Research Report. Princeton, New Jersey: Educational Testing Service.

Mickwitz, L. (2015). *En reformerad lärare: konstruktionen av en professionell och betygsställande lärare i skolpolitik och skolpraktik*. [School reform and the teacher: the construction of a professional and assessing teacher in policy and teacher interviews], Stockholm university.

Ofstedal Telhaug, A., Mediás, O. A., & Aasen, P. (2006). The Nordic model in education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research*, 50(3), 245–283. https://doi.org/10.1080/00313830600743274

Pettersson, A. (1986). *Elevers syn på sin högstadietid och sin räkneförmåga ett år efter avslutad grundskola*. [Students’ view of their lower secondary years and their ability to calculate one year after completing compulsory school], Institutionen för pedagogik, Högskolan för lärarutbildning i Stockholm.

Proches, S. (2016). Descriptive statistics in research and teaching: Are we losing the middle ground? *Quality and Quantity*, 50(5), 2165–2174. https://doi.org/10.1007/s11135-015-0256-3

Prop. 2008/09:6. (2009). *En ny betygsskala*. [A new grading scale], The Swedish Government.

Prop. 2009/10:219. (2010). *Betyg från årskurskurs 6 i grundskolan*. [Grading from grade 6 in compulsory school], The Swedish Government.

Prop. [Government bill] 1992/93:220. (1993). *En ny läroplan: en ny läroplan och ett nytt betygssystem för grundskolan, sameskolan, specialskolan och den obligatoriska särskolan*. [A new curriculum: A new curriculum and grading system for compulsory school, Sami schools, special schools and compulsory school for the intellectually challenged], The Swedish Government.

Räty, H., Kasanen, K., Kiiskinen, J., Nykky, M., & Atjonen, P. (2004). Childrens’ notions of the malleability of their academic ability in the mother tongue and mathematics. *Scandinavian Journal of Educational Research*, 48(4), 413–426. https://doi.org/10.1080/00313830600743365

Rönnberg, L. (2014). Justifying the need for control. Motives for Swedish national school inspection during two governments. *Scandinavian Journal of Educational Research*, 58(4), 385–399. https://doi.org/10.1080/00313831.2012.732605

Salomonsen, H. H., & Andersen, L. B. (2014). Institutional change and professional norms: A study of grading behavior in higher education. *International Journal of Public Administration*, 37(10), 611–624. https://doi.org/10.1080/01900692.2014.903264

Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing*, 18(4), 373–391. https://doi.org/10.1177/026553220101800404

Skolverket. (2001). Utan fullständiga betyg: varför når inte alla elever målen? [Without complete grades: Why do not all pupils reach the goals?], Swedish National Agency for Education.

Skolverket. (2019a). Slutbetyg i grundskolan, våren 2019. [School leaving certificate grades, spring 2019]. PM 2019-09-26 [Memo 26 September 2019]. www.skolverket.se

Skolverket. (2019b). Analyser av likvärdig betygssättning mellan elevgrupper och skolor. [Analyses of equivalent grading between student groups and schools], The Swedish National Agency for Education.

SOU 2016:25. (2016). Likvärdigt, rättssäkert och effektivt: ett nytt nationellt system för kunskapsbedömning: betänkande. [Equivalency, legality and efficiency: a new national system for knowledge assessment: report], Wolters Kluwer.

SOU [Official reports of the Swedish government] 1992:86. (1992). *Ett nytt betygssystem: slutbetänkande*. [A new grading system: Final report], Allmänna förlaget.

Tholin, J. (2006). *Att kunna klara sig i ökänd natur: en studie av betyg och betygskriterier - historiska betingelser och implementering av ett nytt system*. [Coping in infamous nature [sic]: A study of grades and assessment criteria – historical conditions and implementation of a new system], Högskolan i Borås.

Tyrefors Hinnerich, B., & Vlachos, J. (2016). Skillnader i resultat mellan gymnasieelever i fristående och kommunala skolor. [Differences in results between upper-secondary students in independent schools and public schools], Institutet för arbetsmarknadspolitisk utvärdering.

Vlachos, J. (2018). *Trust-based evaluation in a market-oriented school system*. Institutet för näringslivsforskning.

Wikström, C. (2006). Education and assessment in Sweden. *Assessment in Education: Principles, Policy & Practice*, 13(1), 113–128. https://doi.org/10.1080/09695940600563470