



External Evaluation of Schools: Effects of Framework on Results

Carlos Barreira, Graça Bidarra, Maria Piedade Vaz Rebelo

cabarreira@fpce.uc.pt, gbidarra@netcabo.pt, pvaz@mat.uc.pt

Universidade de Coimbra (Portugal)

Abstract

In this paper we will present the framework for external evaluation of schools in Portugal and an analysis of the main results, seeking to highlight some trends.

1. Introduction

Following the publication of Law Nº. 31/2002 on the evaluation of non-higher education, it was initiated, in 2006, a pilot project covering 24 schools (Ministério da Educação, 2006), which is the source of the external evaluation process of schools in Portugal conducted by the Portuguese General Inspectorate of Education (Inspeção-Geral da Educação). This process has gone through several stages and in this academic year (2010-2011) completes the first cycle of external evaluation of schools which began four years ago.

In the framework of reference used by the external school evaluation (ESE) are taken into account five key domains: Results, Provision of Educational Services, School Organization and Management, Leadership, and Adjustment Capacity and School Improvement (Inspeção Geral de Educação, 2009a; Ministério da Educação, 2006). These domains comprise between two and five subdomains, in a total of 19 factors. The ratings of the five domains and factors are assigned in a qualitative scale with four levels: *Very Good*, *Good*, *Sufficient* and *Insufficient*.

The National Board of Education (Conselho Nacional de Educação) has made several opinions on the evaluation reference used in the ESE (Conselho Nacional da Educação, 2008; 2010a, 2010b). In these opinions, it was highlighted the existence of a certain tension between two purposes of evaluation, the formative purpose of school improvement and the summative purpose, the accountability and the responsibility of educational institutions. Thus, it considers that adjustments should be made in the framework of reference of the ESE in order to strengthen the central findings in the domains, and to consider different weights in the domains.

It seems to us, therefore, appropriate to disseminate the main results obtained in this process, presenting a first reading and analysis of published data, trying to show some trends.

2. Methodology

The database of this study was obtained from the results published by the Portuguese General Inspectorate of Education in the Reports on the External Evaluation of School for the academic years 2006-2007, 2007-2008 (Inspeção-Geral da Educação, 2009b), 2008-2009 (Inspeção-Geral da Educação, 2010) and 2009-2010 (Inspeção-Geral da Educação, 2011), where are presented, in percentage terms, the obtained marks in the domains of the framework of reference and the ratings assigned to the factors that constitute them. To ease the data processing, we convert this scale into scores 1, 2, 3 and 4, respectively. The distribution of evaluated schools over this period is presented in Table 1.



Table 1 – Percentages of the classification levels assigned to schools according to the domains of the framework of reference (column-percentages per school year)

Framework of Reference (column percentages per school year)						
Ratings	Academic Results Year	Provision of Educational Services	School Organization and Management	Leadership	Adjustment Capacity and School Improvement	
Very Good	2006-2007	10	14	29	40	11
	2007-2008	4	10	24	32	6
	2008-2009	7	9	23	33	4
	2009-2010	12	13	29	36	4
Good	2006-2007	55	63	61	43	48
	2007-2008	56	59	64	52	37
	2008-2009	60	73	67	51	36
	2009-2010	64	70	63	56	46
Sufficient	2006-2007	34	22	9	16	39
	2007-2008	37	31	11	15	50
	2008-2009	33	18	10	15	54
	2009-2010	24	17	8	8	47
Insufficient	2006-2007	1	1	1	1	2
	2007-2008	3	0	0	1	7
	2008-2009	0	0	0	1	6
	2009-2010	0	0	0	0	3

For each academic year, the score was calculated for each domain by weighting each value of the scale with the percentage with which the domain appears in Table 1 being these values represented in Figure 1. This procedure was also used to analyse the ratings of the factors and their relationship with the domain. Under this analysis, we calculated the average value of the differences between the ratings of each factor and the domain rating. These values are shown in Tables 2 to 6, where the second column represents the value of the rating of the domain and the following columns the difference of rating in the respective factor, presenting the average of the absolute differences in the last row of the tables.

3 Results

3.1 Ratings in the domains

The rating assigned to the different domains that comprise the framework of reference, over the observed academic years (see Table 1 and Figure 1), points to predominance of a positive rating levels of Good / Very Good, with the exception of the domain *Adjustment Capacity and School Improvement*. This trend has not only remained the same, as it has been growing, with the exception of the academic years 2007-2008.

Despite this trend, there are variations depending on the domains. Indeed, the *School Organization and Management* and *Leadership* are those which get a higher percentage of the rating levels Good and Very Good. The domain where these levels are lower is the *Adjustment Capacity and School Improvement*.

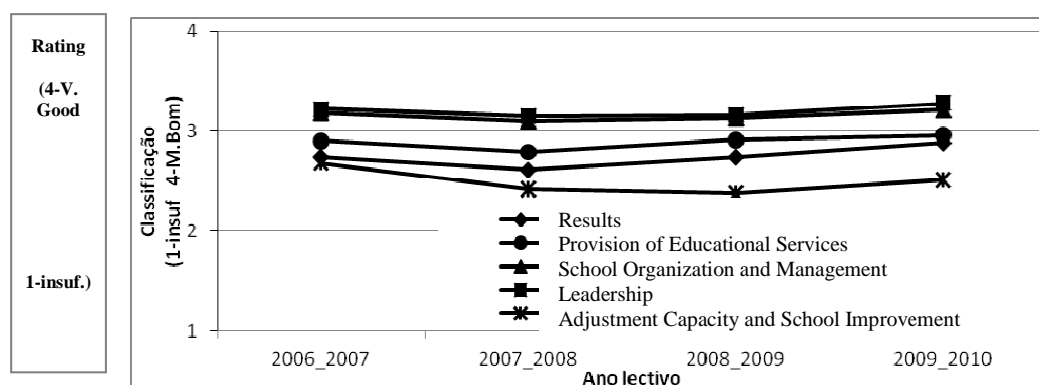


Figure 1 – Evolution of ratings in the domains

It is also possible to verify that the ratings in the domain of *School Organization and Management* are much alike with the ones obtained in the domain of *Leadership*, while the ratings of *Results* are close to the domain of *Provision of Educational Services*.

These data raise the possibility that the interpretation of the goals of School Improvement and Accountability of the model itself may be driving to a linkage between *Results* and *Provision of Educational Services*, with the removal of the rating in relation to the other domains.

3.2 Ratings of the factors

The analysis of the ratings of the factors that characterize the domain *Results* and its evolution, as well as the value of the difference between the rating of each factor and the rating of the domain (see Table 2), shows that the average rating of the *Academic Success* factor has consistently lower values than those obtained in the domain, while the other factors, particularly the factor *Behaviour and Discipline* have a higher average ratings than those obtained in the same domain. That is, the difference of the value between the ratings obtained in the factors and the rating obtained in the domain varies depending on the factors, verifying that the minimum value of that difference corresponds to the factor *Participation and Civic Development*.

Table 2 – Ratings averages obtained in the domain *Results* and the differences between the ratings in the domain and the ratings on each factor.

	Domain	Factors			
Academic Year	Results	Academic Success	Participation and Civic Development	Behaviour and Discipline	Valorisation and Impact of Learning
2006-2007	2.74	-0.18	0.18	0.49	0.20
2007-2008	2.61	-0.25	0.20	0.37	0.19
2008-2009	2.74	-0.21	0.02	0.26	0.16
2009-2010	2.88	-0.26	0.03	0.12	0.16
Average of absolute differences		0.23	0.11	0.31	0.18

When we consider the ratings of the factors included in the domain of *Provision of Educational Services* (see Table 3), it appears that here too there are factors whose average ratings are consistently lower than those of the domain, namely the factors *Coordination and Monitoring* and *Teaching Practice Supervision*, while the factors *Differentiation* and *Supports and Curriculum Support* and *Knowledge* always present higher average ratings than the observed in the domain.



Table 3 – Ratings averages obtained in the domain *Provision of Educational Services* and the differences between the ratings obtained in the domain and the ratings obtained on each factor.

	Domain	Factors			
Academic Year	Provision of Educational Services	Coordination and Monitoring	Teaching Practice Supervision	Differentiation and Supports	Curriculum Support and Knowledge Enhancement
2006-2007	2.9	-0.14	-0.55	0.34	0.19
2007-2008	2.79	-0.27	-0.42	0.39	0.24
2008-2009	2.91	-0.29	-0.61	0.30	0.17
2009-2010	2.96	-0.27	-0.46	0.26	0.26
Average of absolute differences		0.24	0.51	0.32	0.22

Regarding the domain *School Organization and Management* and related factors (see Table 4), we verify that the factors of *Human Resources Management* and *Equity and Justice* obtained ratings slightly higher than those obtained in the domain itself, while factors *Design and Activity Planning* and *Participation of Parents and other members of the Educational Community* obtained ratings well below the ratings obtained in the domain. The factor, which average rating closely matches the rating of the domain is the factor of *Human Resources Management*.

Table 4 – Ratings averages in the domain *School Organization and Management* and differences between the ratings of the domain and the ratings on each factor.

	Domain	Factors				
Academic Year	School Organization and Management	Design/ Activity Planning and Development	Human Resources Management	Financial and Material Resources Management	Participation of Parents and other members of the Educational Community	Equity and Justice
2006-2007	3.18	-0.18	0.01	0.02	-0.35	0.08
2007-2008	3.10	-0.24	0.06	0.00	-0.31	0.12
2008-2009	3.13	-0.23	0.00	-0.13	-0.25	0.07
2009-2010	3.21	-0.27	0.11	-0.04	-0.19	0.06
Average of absolute differences		0.23	0.04	0.05	0.28	0.08

In the domain of *Leadership* (see Table 5), the factors *Motivation and commitment* and *Partnerships, protocols and projects* obtained slightly higher ratings than those obtained in the domain itself, while factors *Vision and strategy* and *Innovation opening* present average ratings lower than those obtained in the domain. The rating of the factor *Motivation and commitment* is the factor that matches closely to the rating obtained in the domain.

Table 5 – Ratings averages obtained in the domain *Leadership* and differences between ratings of the domain and the ratings obtained on each factor.

	Domain	Factors			
Academic Year	Leadership	Vision and strategy	Motivation and commitment	Innovation	Partnerships, protocols and projects
2006-2007	3.22	-0.12	0.00	-0.09	0.07
2007-2008	3.15	-0.19	0.12	-0.22	0.05
2008-2009	3.16	-0.22	0.09	-0.20	0.17
2009-2010	3.28	-0.22	0.09	-0.36	0.12
Average of absolute differences		0.19	0.08	0.22	0.10

Finally, with regard to the factors that constitute the domain *Adjustment Capacity and School Improvement* (see Table 6), we observed both in the domain and in the two factors there are average ratings below 2.5, and the ratings obtained in the factors are close to those obtained in the domain.



Table 6 – Ratings averages obtained in the domain of *Adjustment Capacity and School Improvement* and differences between the ratings in the domain and ratings on each factor.

	Domain	Factors	
Academic Year	Adjustment Capacity and School Improvement	Self-Evaluation	Sustainability and progress
2006-2007	2.68	-0.11	0.02
2007-2008	2.42	-0.08	0.05
2008-2009	2.38	-0.05	0.00
2009-2010	2.51	-0.03	-0.13
Average of absolute differences		0.07	0.05

4. Discussion and conclusion

The data presented show that in general the schools have achieved good ratings in all domains, except on the *Adjustment Capacity and School Improvement*. We also observed that there are domains whose ratings are consistently higher and very close, namely, *Leadership* and *School Organization and Management*, being the ratings of the domains of *Results* and *Provision of Educational Services* consistently close in numbers. However, these data also suggest the existence of a differential valuing of the domains in accordance with the purposes of the model.

The observation of a certain regularity of the ratings over the years under review, and the fact that the contribution of different factors for the domain ranking varies between them, suggests that, although there is no weighting factors in the framework of reference, and defending that it should exist, it is somehow inherent in the evaluation process, when we verify the ups and downs of the ratings in the factors related to the corresponding domain.

The presented data even allow raising questions about the conceptualization of their own domains and the factors that characterize them. These aspects appeared to us to be of much interest and relevance for further study, which is already underway, on a framework of reference and the effects on evaluation related to ratings in the domains and the factors of the schools under evaluation.

5. Bibliography

- [1] Conselho Nacional de Educação (2008). Parecer n.º 5, Diário da República 2ª Série – N.º 113 – de 13 de Junho.
- [2] Conselho Nacional de Educação (2010a). Parecer n.º 3, Diário da República, 2ª Série – N.º 111 – 9 de Junho.
- [3] Conselho Nacional de Educação (2010b). Recomendação sobre avaliação das escolas (relatores Jorge Miguel Marques da Silva e Paula Santos). Lisboa: CNE.
- [4] Inspeção-Geral da Educação (2008). Avaliação externa de escolas - Relatório 2006-2007. Lisboa: Ministério da Educação - IGE.
- [5] Inspeção-Geral da Educação (2009a). Avaliação externa de escolas - Referentes e instrumentos de trabalho. Lisboa: Ministério da Educação - IGE.
- [6] Inspeção-Geral da Educação (2009b). Avaliação externa de escolas - Relatório 2007-2008. Lisboa: Ministério da Educação - IGE.
- [7] Inspeção-Geral da Educação (2010). Avaliação externa de escolas - Relatório 2008-2009. Lisboa: Ministério da Educação - IGE.
- [8] Inspeção-Geral da Educação (2011). Avaliação externa de escolas - Relatório 2009-2010. Lisboa: Ministério da Educação - IGE.
- [9] Lei n.º 31/2002. Diário da República – I Série - A N.º 294 de 20 de Dezembro (aprova o sistema de avaliação da educação e do ensino não superior).
- [10] Ministério da Educação (2006). Relatório final da actividade do Grupo de Trabalho para Avaliação das Escolas. Lisboa: CNE.