# Evidence based decision, data science and school dropout: a complexity approach

Marcos Cavalcanti[1]

**Abstract**

The use of data and evidence for decision-making seems an obvious choice. We do not expect decision-making to be based on personal opinions or "hunches". When dealing with problems that we do not know or have no opinion about, we are more open to using data and evidence to support our decisions. However, if we have an opinion on the subject studied, our decision is contaminated, and we tend not to consider data and evidence. School dropout is one of these throbbing subjects, which arouses a lot of interest from researchers, but where we often use data and evidence only to corroborate preconceived opinions. The subject has been approached by several areas of knowledge to understand the factors that involve the possible causes that lead to abandonment. Despite the various explanations and causes for school dropout, everyone seems to agree that this is a relevant and urgent problem. These factors are analyzed in the literature in an independent manner and often present an additional problem: because these are factors that cannot or a difficult to be changed (such as race, ethnicity or gender) and requires a long-term for being changed (such as inequality, poverty and racism), improving the situation of the students who are currently at school is difficult if not impossible.

**Keywords:** School dropout. Education. Data Science. Evidence-based decision. Complexity.

[1] Coordenador do CRIE/UFRJ e professor da UFRJ. E-mail: marcos@crie.ufrj.br

## 1. INTRODUÇÃO

School dropout is usually defined as “leaving education without obtaining a minimal credential, most often a higher secondary education diploma" (De Witte et al 2013, p.14).

Given the relevance of the theme, several government programs have been developed to reduce school dropout. The United States created in 2001 the program ''No Child Left Behind[2]'' (2001), while the European Community created the programs '' Lisbon 20003 '', and more recently the program '' Europe 20204 ''. Despite increasing attention from policy makers, none of these programs has managed to significantly reduce school dropout rates. In addition, with the pandemic, the problem of school dropout brought new challenges in the construction of educational policies worldwide. One of the reasons for this has to do with a common view shared by these programs and studies on this topic of seeking one single cause for abandonment or, at best, a limited set of reasons for it. Few articles and programs have applied a systemic and complex approach to understand and describe the phenomena, seeking to identify the connections between the various factors and their dependence on the different contexts where abandonment occurs.

The use of data and evidence for decision-making seems an obvious choice. We do not expect decision-making to be based on personal opinions or “hunches”. When dealing with problems that we do not know or have no opinion about, we are more open to using data and evidence to support our decisions. However, if we have an opinion on the subject studied, our decision is contaminated, and we tend not to consider data and evidence. Perhaps unconsciously, we tend to privilege our opinions and seek only data that prove our beliefs, even if the evidence points out to the opposite direction (KVERNBEKK, 2017).

School dropout is one of these throbbing subjects, which arouses a lot of interest from researchers, but where we often use data and evidence only to corroborate preconceived

[2] https://www2.ed.gov/nclb/landing.jhtml
[3] https://www.europarl.europa.eu/summits/lis1_en.htm
[4] https://www.education.gouv.fr/media/72354/download

opinions. The subject has been approached by several areas of knowledge to understand the factors that involve the possible causes that lead to abandonment.

To analyze how academic research has investigated this problem, we explored a bibliographic survey in the open access databases. With this survey, we identify the causes pointed out in the literature for school dropout using a search tool for academic articles and books using generically the keyword "school dropout".

In the Google Scholar database, in the period from 2010 to 2020 (09/07/2020), there were approximately 1,430 articles with the keyword in the article title and 24,000 articles that used it anywhere in the article. In order to improve the research, the combination “School dropout” and “high school” was used to focus on school dropouts, which is the focus of our research, and thus we found 478 articles with these words in the title of the article. Finally, the keyword “evidence” was added. Incorporating this keyword to our search was crucial, as our study is synthetized as follows: making decisions about school dropout based on data and evidence. The search resulted in 16 articles.

The same searching approach was replicated in the Scopus Elsevier database, considering the same period from 2010 to 2020 (09/07/2020). It found 9 articles, 4 of which also happen to be in the Google Scholar search.

Table 1 - Base Scopus Results

| Authors | Title | Year | Cited by |
|---|---|---|---|
| Butterworth P., Leach L.S. | Early Onset of Distress Disorders and High-School Dropout: Prospective Evidence from a National Cohort of Australian Adolescents | 2018 | 6 |
| Migali G., Zucchelli E. | Personality traits, forgone health care and high school dropout: Evidence from US adolescents | 2017 | 4 |
| Cabus S.J., De Witte K. | Why Do Students Leave Education Early? Theory and Evidence on High School Dropout Rates | 2016 | 11 |
| Barth J.R., Cebula R.J., Shen I.-L. | Is the high school dropout rate an increasing function of the proportion of | 2016 | |

| | | | |
|---|---|---|---|
| | the population in the US cities that is Hispanic? Exploratory evidence | | |
| Mo D., Zhang L., Yi H., Luo R., Rozelle S., Brinton C. | School Dropouts and Conditional Cash Transfers: Evidence from a Randomized Controlled Trial in Rural China's Junior High Schools | 2013 | 30 |
| Crofton S.O., Anderson W.L., Rawe E.C. | Do higher real minimum wages lead to more high school dropouts? Evidence from Maryland across races, 1993-2004 | 2009 | 6 |
| Temple J.A., Reynolds A.J., Miedel W.T. | Can early intervention prevent high school dropout? Evidence from the Chicago child-parent centers | 2000 | 49 |
| Rees D.I., Mocan H.N. | Labor market conditions and the high school dropout rate: Evidence from New York State | 1997 | 30 |
| Rumberger R.W. | High School Dropouts: A Review of Issues and Evidence | 1987 | 474 |

So, it was decided to analyze the articles presented on Google Scholar and Scopus as a starting point for the exploratory research. An analysis of these articles revealed several reasons why young people drop out of school. These reasons can be grouped into two main categories:

**A)** Factors related to the student and the family: such as adolescent pregnancy, family income, gender, parental educational level, parental employment, race / ethnicity - blacks and Latinos are more likely to dropping out of school[5], social class or socioeconomic status, student academic performance. (Ekstrom, 1986; Rumberger, 1987; Temple et al., 2000; Entwisle et al. 2004; Allensworth, 2004, 2005; Duchesne et al, 2005; Ishitani, Snider, 2006; Koball, 2007; Dustmann, van Soest, 2008; Crofton et al. 2009; Dalton et al. 2009; Barth et al. 2016; Migali, Zucchelli, 2017; Butterworth, Leach, 2018)

**B)** Factors related to the school and the community (context): class size, absence of playgrounds and green areas, neighborhood characteristics (geographic location of families' residence, possible housing problems, poverty, school resources, teacher-student ratio, type of school (public, private, more or less selective), unemployment rates, and violence. (Rumberger,

5

1983, 2004a ; Ekstrom, 1986; Pitman, 1993; Rees, Mocan, 1997; Kaufman et al. 2004; Balfanz, Legters, 2004; Blue, Cook, 2004; Dustmann, van Soest, 2008; Crofton et al., 2009; Mo Di et al, 2013; Cabus, De White, 2016).

In sum, according to a forecited literature, the reasons for school dropout have focused on factors that are difficult to sort out in a short and medium term.

Despite the various explanations and causes for school dropout, everyone seems to agree that this is a relevant and urgent problem.

According to a study "The consequences of dropping out of high school" (Sum et al, 2009), the average student who drops out of school can expect to earn an annual income of US$20,241, according to the US Census Bureau. That is $ 10,386 less than a typical high school graduate earns and $36,424 less than someone with a college degree under the same period. And the number of people who drop out of school, although it has fallen in recent years, is still high in the USA. There were 2.1 million status dropouts between the ages of 16 and 24 in 2018, and the overall status dropout rate was 5.3 percent. In 2014, the rate was 6.5% of a school population of 50,468,456 students.

According to Sum et al (2009), half of Americans who need public assistance are dropouts. This study found that each dropout from high school costs American taxpayers $292,000, or just over 957 billion per year, 4.65% of the USA GDP (Gross Domestic Product) of 20,540 trillion dollars.

A similar study produced by the Fernandes,R (2018) supported by Unibanco Institute and applied to Brazil, showed that 575,000 young people did not complete elementary school in 2018. According to the study, each dropout before completing high school costs $67,636 (sixty-seven thousand dollars) per young person, or $38 billion per year, one loss equivalent to 3% of Brazilian annual GDP, of $ 1.1 trillion in 2018.

In addition to enormous personal costs (such as low self-esteem, greater difficulties in finding well-paid work, greater tendency to get involved with crime and violence and die earlier), school dropout means a significant economic loss for a country. Addressing this problem therefore means not only an individual gain for those involved, but a real and concrete benefit for the whole of society.

**A critical approach: it is not complicated, it is complex…**

"For every complex problem, there is a clear, simple and wrong answer." H.L.Mencken

As noted, the school dropout literature indicates that the reasons for dropping out of school are often associated with socioeconomic aspects such as long-term unemployment (from parents), poverty, race / ethnicity, sustained dependence on public assistance, single parenthood (in women), political and social apathy and crime. (Christenson et al., 2001; Rumberger, Land, 2003; Kaufman et al. 2004; Vitzcain, 2005).

Most of the articles we reviewed attempt to identify the predictive variables, those factors that determine abandonment, and the government programs sought to take actions to deal with these various variables and factors. But changing socio-economic conditions (poverty, inequality, long-term unemployment) requires a long period of time and does not affect students who are already in school. Likewise, pointing out reasons of ethnicity / race / gender does not help to find concrete solutions for students who are already in school. After all, you cannot change a society's structural racism in the short-term change neither students' skin color or ethnic origin.

In fact, in line with a decision-making paradigm based on data and evidence, I defend the idea that we need to change our mindset and adopt a complex approach to assess the problem. In My point of view, the issue of clarity and education is not a complicated problem. It is complex.

Glouberman and Zimmerman (2002) provide an excellent illustration of the differences between simple, complicated, and complex problems. In a simple system, like using a recipe to make an apple pie, even inexperienced people can follow the recipe and make the pie, obtaining satisfactory results. In addition, we managed to produce approximately the same results in different situations, with different people.

In complicated problems, like sending a rocket into space, we need several high-level specialists in various fields of knowledge to generate a successful result. However, once that result is achieved it is, in most cases, replicable.

According to Glouberman and Zimmerman (2002), in simple contexts we can easily correlate cause and effects. This is the realm of the known. Situations can be clearly defined, and appropriate responses identified. The policy maker's role is to delegate, use best practices and communicate standard operating procedures to be followed clearly and directly. Complicated contexts are the realm of experience and data analysis. We know what we do not know. Cause and effect are not evident, but they can be discovered through analysis. The role of the policy maker here is to bring together the necessary minds and skills and to encourage differing opinions, to seek answers to what we do not know. We usually think of educational problems this way...

Over time, potential predictors of non-graduation have generally been sought: first, among students and their families, later in schools, teachers and colleagues, and only at a later stage in the broader context or environment (Rumberger, 2004a). In addition, attention is often placed on immutable factors (socioeconomic factors and other intractable risk factors, such as gender, race and ethnicity, parenting or mother tongue), creating the impression that early school leaving is partly a natural process, which depends fundamentally on these factors.

The ‘‘abandonment discourse’’ has thus associated early school leaving with unemployment, urban poverty, and juvenile delinquency (often serving as a substitute for race and class) (Dorn, 1996). The stereotype, par excellence, of the adolescent man “devoid of culture”, unintelligent, unqualified, not adjusted, not white, who ends up unemployed and delinquent, has been increasingly used by literature. Some studies, however, have shown the weakness of the arguments of those who naturalized school dropout. Herbert and Reis (1999), showed that most young people who present the factors considered determinants to explain dropout (poverty, unemployment, ethnicity / race, income, etc.) do not drop out of school. Moreover, some of them performed very well.

Dropout should not be a direct consequence of the environment, inequality or the person's ethnicity / race. While recognizing that this is a multi-dimensional problem, the way the articles address the problem is based on a complicated problem approach: a straightforward, not complex, cause-solution approach. The causes are almost all related to this socioeconomic origin,

to these immutable factors. And, they make a strong correlation between evasion and defaults and unemployment.

If we follow these analyzes, there is not much to be done, as the skin color cannot be changed, structural racism will take time to end and inequality will last for a few years before being reduced.

The issue of school dropout and education must be approached with a complex view. The complex is the realm of unknown unknowns. We do not know what we do not know. It is a space of constant change and unpredictability. There are no right answers, only emergent behaviors, a concept discussed later in this article. The role of the policy maker in these situations is to create spaces for interaction and experimentation for the emergence of patterns, which is best done by increasing the levels of interaction and communication within the system. Experience is useful, but not enough to solve complex problems. We need to test and experiment, certainly, but we need to have patience and a watchful eye for new patterns of behavior.

The complex nature of educational governance, involving countless layers and actors, can be an overwhelming problem, without a clear input for policymakers. Traditional approaches, which generally focus on proposals for solving complicated problems, have proved ineffective. For all we know about complex problems, programs and standard initiatives, which must be replicated in all places and situations, they are excellent for sending a rocket into space, but they do not serve to deal effectively with rapidly evolving and expanding ecosystems that are modern education systems.

Educational initiatives usually work in the realm of the complicated when they are operating in the realm of the complex (Duit et al., 2010). Iterative feedback is often limited in this approach, and flexibility is not always a high priority in the design of the initiative. They forget that complex problems cannot be adequately captured using such linear approaches (Morrison, 2010; Duit et al. 2010).

The premise of such linear models is that inputs to the system will result in predictable results. Although appropriately predictive of some static and closed systems, these models fail to adequately predict behavior or capture the essence and emerging properties of complex systems involving three or more components of interaction (Johnson, 2008).

What works for one child, teacher, district, or system is not guaranteed to work for another. This makes the problem of educational governance more complex than complicated, as the solutions are not necessarily replicable and transferable.

The bibliographic review carried out in this article shows that the main causes pointed out in the literature can be grouped into two major groups: Factors related to the student and the family (family income, gender, parental educational level, race and ethnicity, social class or socioeconomic status) and Factors related to the school and the community – context (class size, neighborhood characteristics, resources invested in schools, teacher-student ratio, type of school, unemployment rates and violence).

These factors are analyzed in the literature in an independent manner and often present an additional problem: because these are factors that cannot or a difficult to be changed (such as race, ethnicity or gender) and requires a long-term for being changed (such as inequality, poverty and racism), improving the situation of the students who are currently at school is difficult if not impossible.

The ''discourse of abandonment'' has thus associated early school leaving with unemployment, urban poverty and juvenile delinquency, generating a stereotype, par excellence, of adolescents as a being "devoid of culture", unintelligent, unqualified , not adjusted, not white, who ends up unemployed and delinquent. Some studies (Herbert and Reis, 1999) have already shown, however, that most young people who have these factors do not drop out of school. Some of them performed very well actually.

Some studies show that school delay / lag is, in fact, one of the factors that contribute to dropout6. Our study shows that the problem has its origin in a deficient literacy process. Poor literacy will prevent this student from being able to continue with peers of the same age and suggests that the effort to reduce the dropout rate should focus on literacy.

[6] See A critical review of the literature on school dropout (2013) Kristof De Witte a, c, ⇑, Sofie Cabus a, Geert Thyssen a, b, Wim Groot a, d, Henriëtte Maassen van den Brink a, d)


DOI: 10.36559/v4512022

In this article, I show that the school dropout problem is not a complicated one. From my point of view, education is a complex problem, and we should make decisions based not on what we “think”, but on data and evidence.

**REFERENCES**

Allensworth, E. M., 2004. Ending social promotion: Dropout rates in Chicago after implementation of the eight-grade promotion gate. Consortium on Chicago School Research. https://consortium.uchicago.edu/publications/ending-social-promotion-dropout-rates-chicago-after-implementation-eighth-grade (accessed June 16 2020)

Allensworth, E. M., 2005. Dropout rates after high stakes testing in elementary school: A study of the contradictory effects of Chicago’s efforts to end social promotion. Educational Evaluation and Policy Analysis, 27(4), 341–364. https://doi.org/10.3102/01623737027004341

Balfanz, R., Legters, N., 2004. Locating the Dropout Crisis. Which High Schools Produce the Nation's Dropouts? Where Are They Located? Who Attends Them? Report 70. Center for Research on the Education of Students Placed at Risk CRESPAR. https://eric.ed.gov/?id=ED484525 (accessed July 14 2020)

Barth, J. R., Cebula, R. J., Shen, I. L. 2016. Is the high school dropout rate an increasing function of the proportion of the population in the US cities that is Hispanic? Exploratory evidence. Applied Economics Letters, 23(15), 1099-1103. https://doi.org/10.1080/13504851.2015.1136391

Blue, D., Cook, J. E., 2004. High school dropouts: Can we reverse the stagnation in school graduation. Study of High School Restructuring, 1(2), 1-11.

Butterworth, P., Leach, L. S., 2018. Early onset of distress disorders and high-school dropout: Prospective evidence from a national cohort of Australian adolescents. American journal of epidemiology, 187(6), 1192-1198 https://doi.org/10.1093/aje/kwx353

Cabus, S. J., De Witte, K. 2016. Why do students leave education early? Theory and evidence on high school dropout rates. Journal of Forecasting, 35(8), 690-702. https://doi.org/10.1002/for.2394

Christenson, S. L., Sinclair, M. F., Lehr, C. A., & Godber, Y. (2001). Promoting successful school completion: Critical conceptual and methodological guidelines. School Psychology Quarterly, 16(4), 468–484. https://doi.org/10.1521/scpq.16.4.468.19898

Crofton, S., Anderson, W., Rawe, E., 2009. Do Higher Real Minimum Wages Lead to More High School Dropouts? Evidence from Maryland across Races, 1993-2004. The American Journal of Economics and Sociology, 68(2), 445-464. http://www.jstor.org/stable/27739779

Dalton, B., Glennie, E., Ingels, S. J. 2009. Late High School Dropouts: Characteristics, Experiences, and Changes Across Cohorts. Descriptive Analysis Report. NCES 2009-307. National Center for Education Statistics. https://eric.ed.gov/?id=ED505580 (accessed June 20 2020)

De Witte, K., Cabus, S., Thyssen, G., Groot, W., van Den Brink, H. M., 2013. A critical review of the literature on school dropout. Educational Research Review, 10, 13-28. DOI: https://doi.org/10.1016/j.edurev.2013.05.002

Dorn, S., 1996. Creating the dropout. An institutional and social history of school failure. Westport, CT/London: Praeger. https://eric.ed.gov/?id=ED399617 (accessed June 25 2020)

Duit et al, 2010, Governance, complexity, and resilience, Global Environmental Change, volume 20, issue 3, Aug 2010, pages 363-368

Dustmann C., van Soest A. 2008 Part-time work, school success and school leaving. In: Dustmann C., Fitzenberger B., Machin S. (eds) The Economics of Education and Training. Studies in Empirical Economics. Physica-Verlag HD. 23-45 https://doi.org/10.1007/978-3-7908-2022-5_3

Duchesne, S. et al. 2005. 'The transition from elementary to high school: The pivotal role of mother and child characteristics in explaining trajectories of academic functioning', International Journal of Behavioral Development, 29(5), pp. 409–417. https://doi.org/10.1177/01650250500206067

Ekstrom, R. B., 1986. Who drops out of high school and why? Findings from a national study. Natriello, Gary (ed.) School Dropouts: Patterns and Policies. New York: Teachers College Press. p 52-69

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2004). Temporary as compared to permanent high school dropout. Social Forces, Volume 82 (3), 1181 - 1205. https://doi.org/10.1353/sof.2004.0036

Fernandes, R. 2010. Ensino Médio: Como aumentar a atratividade e evitar a evasão. Instituto Unibanco. p. 44 https://www.institutounibanco.org.br/wp-content/uploads/2013/07/ensino_medio-como_aumentar_a_atratividade_e_evitar_a_evasao.pdf (accessed May 15 2020)

Glouberman, S., Zimmerman, B., 2002. Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like? Commission on the Future of Health Care in Canada. p. 37 http://publications.gc.ca/pub?id=9.558204&sl=0 (accessed Apr 10 2020)

Herbert, T. P., Reis, S. M. 1999. Culturally diverse high-achieving students in an urban high school. Urban Education, 34(4), 428–457. https://doi.org/10.1177/0042085999344002

Ishitani, T. T., Snider, K. G. 2006. Longitudinal effects of college preparation programs on college retention. IR Applications, 9, 1–10. https://eric.ed.gov/?id=ED504377 (accessed June 16 2020

Johnson E. S., 2008. Ecological Systems and Complexity Theory: Toward an Alternative Model of Accountability in Education, International Journal of Complexity and Education, vol 5, no 1, https://doi.org/10.29173/cmplct8777

Kaufman, P., Alt, M. N., Chapman, C. D., 2004. Dropout Rates in the United States: 2001. Statistical Analysis Report National Center for Education. p.94 https://nces.ed.gov/pubs2005/dropout2001/ (accessed May 05 2020)

Koball, H. 2007. Living arrangements and school dropout among minor mothers following welfare reform. Social Science Quarterly, 88(5), 1374-1391. https://doi.org/10.1111/j.1540-6237.2007.00507.x

Kvernbekk, T. 2017. Evidence-Based Educational Practice. Oxford Research Encyclopedia of Education. Retrieved 3 Oct. 2020, from https://doi.org/10.1093/acrefore/9780190264093.013.187

Mandinach, E. B., Schildkamp, K. 2020. Misconceptions about data-based decision making in education: An exploration of the literature. Studies in Educational Evaluation https://doi.org/10.1016/j.stueduc.2020.100842

Migali, G., Zucchelli, E., 2017. Personality traits, forgone health care and high school dropout: Evidence from US adolescents. Journal of Economic Psychology, 62, 98-119. http://dx.doi.org/10.1016/j.joep.2017.06.007

Mo, D., Zhang, L., Yi, H., Luo, R., Rozelle, S., Brinton, C., 2013. School dropouts and conditional cash transfers: evidence from a randomized controlled trial in rural China's junior high schools. The Journal of Development Studies, 49(2), 190-207. https://doi.org/10.1080/00220388.2012.724166

Morrison, K. 2010 Complexity Theory, School Leadership and Management: questions for Theory and Practice, Educational Management Administration & Leadership, volume 38, issue 3.

Pittman, R. B. 1993. The 21st century and secondary school at-risk students: What is ahead for teachers in rural America? (Conference Proceedings). https://eric.ed.gov/?id=ED359003 (accessed Apr 22 2020)

Sum, A., Khatiwada, I., McLaughlin, J., Palma, S. 2009. The consequences of dropping out of high school. Center for Labor Market Studies Publications, 23

Rees, Daniel I. & Mocan, H. Naci, 1997. Labor market conditions and the high school dropout rate: Evidence from New York State, Economics of Education Review, Elsevier, vol. 16(2), 103-109.

Rumberger, R. W. 1987 High school dropouts: A review of issues and evidence. Review of educational research, v. 57, n. 2, 101-121. https://doi.org/10.3102/00346543057002101

Rumberger, R. W. (1983). Dropping out of high school: The influence of race, sex, and family background. American Educational Research Journal, 20(2), 199-220. https://doi.org/10.2307/1162594

Rumberger, R. W., Lamb, S. P. 2003. The early employment and further education experiences of high school dropouts: A comparative study of the United States and Australia. Economics of Education Review, 22 (4), 353–366. https://doi.org/10.1016/S0272-7757(02)00038-9

Rumberger, R. W. (2004a). Why students drop out of school? In G. Orfied (Ed.), Dropouts in America: Confronting the graduation rate crisis, Cambridge. 131–155. MA: Harvard Education Press.

TEMPLE, J.A.; REYNOLDS, A. J.; MIEDEL, W.T. 2000. Can early intervention prevent high school dropout? Evidence from the Chicago Child-Parent Centers. Urban education, v. 35 (1), 31-56 https://doi.org/10.1177/0042085900351003

Vizcain, D. C., 2005 Investigating the Hispanic/Latino Male Dropout Phenomenon: Using Logistic Regression and Survival Analysis. Graduate Theses and Dissertations. https://scholarcommons.usf.edu/etd/896