

行政院國家科學委員會補助專題研究計畫成果報告

教育爭議的衍生：一個行為決策學的分析

計畫類別： 個別型計畫

計畫編號： NSC 94-2413-H-039-002-

執行期間： 94年08月01日至95年07月31日

計畫主持人： 黃榮村

計畫參與人員： 王傳華（助理）

成果報告類型(依經費核定清單規定繳交)： 精簡報告

本成果報告包括以下應繳交之附件： N/A

執行單位： 中國醫藥大學

中 華 民 國 95 年 12 月 14 日

教育爭議的衍生：一個行為決策學的分析

摘要

台灣在 2002-2004 年引發十年教改之大規模爭議，本研究蒐集 595 篇評論文章與研究論文，從行為決策角度分析三類主要發生在 2002-2004 年間的群聚論證，並以基礎比例忽略 (base-rate neglect)、九年一貫課程改革、教育爭議之局部均衡等項，作細部分析。本研究並討論政策因應在多重判準下之共識尋找中所擔負的角色，並提出冷卻策略雖非實質共識尋找下之最主要成分，但有助於啟動理性過程以獲致教育理念與實踐上的真正均衡。

關鍵詞：教育爭議、行為決策、基礎比例忽略、課程改革、共識尋找、局部均衡

The Emergence of Educational Controversies: **A Behavioral Decision Analysis**

Abstract

A large-scale rebuttal on the “Ten-Year Taiwan Education Reform” was prevailing in the year 2002-2004. This study is designed to collect and classify the well-defined arguments along a framework of behavioral decision analysis. 595 articles were selected with a chronicle distribution that showed a natural central tendency around the year 2002-2004. Three clusters of educational debates are identified and put under an analytical framework of behavioral decision theory. Each cluster is further explained with its association to the study of base-rate neglect, a national debate on the curriculum reform, and local equilibrium for the educational controversies. Finally, the study discusses the role of policy response on the consensus seeking under multiple criteria. It is proposed that although a cooling-down practice is not the substantial part of real consensus seeking, it is still a prerequisite for initiating a rational approach to find the genuine equilibrium.

Keywords: educational controversy, behavioral decision, base-rate neglect, curriculum reform, consensus seeking, local equilibrium

The Emergence of Educational Controversies: **A Behavioral Decision Analysis**

Introduction

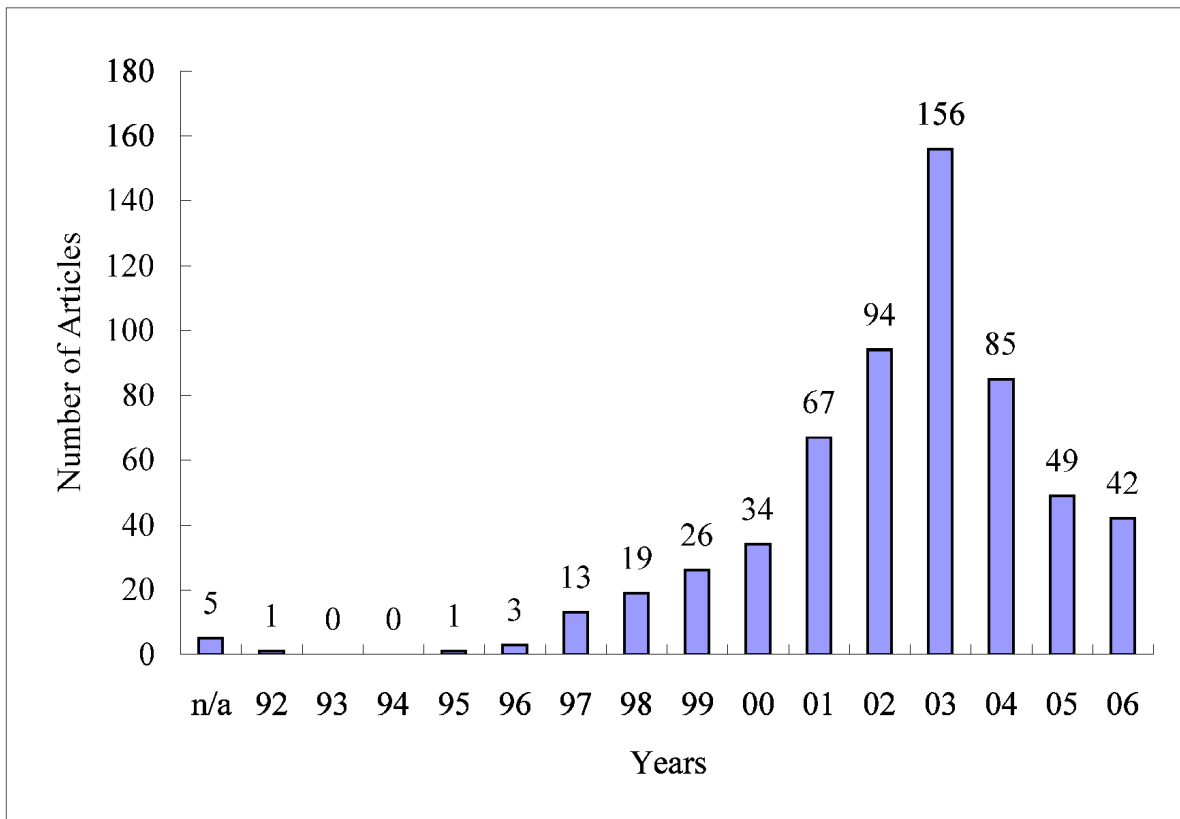
Educational reforms in Japan that prevailed during the late 1980s and 1990s emphasize a lift of stress on individuality and encouraging creativity by incorporating more freedom and choice into the education system. The tone manifested in Japan society was basically a liberal agenda embedded in a free market mechanism. The approach was later bitterly criticized and the main reform measures to date favor a progressive rather than a neoliberal direction (Cave, 2001). Put it more clearly, it seems to be a rebound from liberal to conservative in a systematic way to alleviate the societal unrest that was incurred through the so-called “harsh process” of educational reform in Japan.

Similar observations can also be vividly witnessed in Taiwan. A large-scale rebuttal on the coined “Ten-Year Taiwan Education Reform” (Pan & Yu, 1999) was prevailing in the year 2002-2004. The reason is obvious in that most parents are anxious on whether their children were educated properly in the schools and almost every citizen concerns with the proper cultivation of next-generation talents. The outbreak of systematic rebuttals might also be reinforced by the accumulating tension that arose from the political transition during this period. A widened difference in political ideology after the regime turnover was waiting for a common-interest topic to initiate a cross-fire in the battlefield of approaching Presidential election campaign in 2004. Overwhelming discontents with educational quality could thus easily find a way to blame the past doings of the Ten-Year Education Reform.

There are at least two contributing factors that are responsible for this emergence of educational controversies. One is due to the inefficiency of management and a failure to provide sufficient support for dissolving the difficulties in encountering with accumulated deficiencies and in implementing appropriate education reform projects. The other arises from an entanglement of different conceptual extremes along the multidimensional educational continuum. Some arguments in the controversy are legitimate and well-informed in their own rights, nonetheless, some are unwarranted assertions with slim data support. This study intends to collect and classify the well-defined arguments along a framework of behavioral decision analysis in the hope that the popular fallibility which has been reliably demonstrated in human judgment can also be identified. The difficulty also lies in the fact that different educational claims can not be successfully accommodated. The study will henceforth discuss the possible modifications that might help achieve a local equilibrium on some cases.

Classification of Educational Controversies

A longitudinal data collection that spanned largely through 2001 to 2006 was conducted on the following four major sources: (1) the data bank of the Ministry of Education, especially from the National Educational Archives; (2) electronic local educational journals; (3) educational websites; and (4) news selection of www.chinatimes.com. More than 841 articles were inspected and a total of 595 were finally selected. Figure 1 depicts a chronicle representation of the selected 595 articles.



**Figure 1 A chronicle distribution of 595 articles
(n/a: unable to identify the publication date).**

The selected writings were then combined to extract 46 subtopics. Each subtopic was further tabulated with tallies that appeared on all writings. There are a total of 1580 tallies, therefore, each subtopic was cited on average 35 times across 595 articles. Three major clusterings can then be identified:

1. 200 citations were centered around the discussion of joint college entrance examination and a rapid growth of universities and colleges.
2. 484 citations focused on the issues of unacceptable equity and fairness in learning or on high-school entrance examination. The discussion of adverse effect of newly-launched

curriculum reform is popular in these citations.

3. 876 citations criticized the discrepancy and its unintended consequences between claimed ideals and practiced reality in educational reforms. The debate was also urging the adoption of a progressive and conservative educational and curricular reform, a centralized edition of textbook in the primary and secondary education, and even a claim to stop the undergoing reformed curriculum. A rethinking of tracking and ability grouping is also classified under this category.

The first cluster is obviously on tertiary education. The second and third clusters deal with primary and secondary education, especially on the junior high school system. The three clusters cover a total of 1560 citations which accounts for 99% share of 1580 tallies.

A Behavioral Decision Analyses for Three Types of Educational Debates

A closer look of the above-mentioned clusters through a behavioral decision analysis will justify the following correspondences. Cluster 1 may correspond to certainty events that need clear and accurate data to get convinced. The study suggests a notion of base-rate neglect to accommodate these differences. Cluster 2 corresponds to risky events that need a convincing data set to clarify the unwarranted assertions. The study will discuss the debate on curriculum reform with the accompanying worries over the worsening of educational achievement. Cluster 3 corresponds to uncertain events with unknown probability structure and outcomes. It needs a consensus under multiple criteria, henceforth, an equilibrium problem.

The nature of Clusters 3 is more emotional toned than in Cluster 2. Hsu *et al.* (2005) found that emotional factors are easier to intrude in the processing of ambiguity rather than in risk judgments. A subjective feeling of justice and fairness is easily aroused in this line of reasoning or in the debate of Cluster 3 issues. Human beings are tempting to act as an intuitive prosecutor (Tetlock, 2002), usually accompanied with strong emotional reaction to urge regulatory measures to restore order or justice (Fehr & Fischbacher, 2003; de Queriain *et al.*, 2004). That will make the debate even more difficult to achieve equilibrium. Following this line of discussion, we can classify Cluster 1 as “easy” problem, Cluster 2 as “not-so-easy” problem, and Cluster 3 the “hard” problem.

Base-Rate Neglect

A common theme of Cluster 1 debate is the over-production of college students and, henceforth, as serious decline of educational quality. The data for arguments is a popular assertion of 90% or even higher admission rate to the universities among 18-22 years old population. The assertion is unwarranted anyway. The educational statistics of the year 2006 showed that (1) gross enrollment rate of tertiary education for 18-22 population > 80%, (2) net enrollment rate of tertiary education for 18-22 population = 53.51% (3) admission rate to the universities and colleges through paper-pencil joint entrance examination = 85~90%, and

(4) net enrollment rate to the universities and colleges for 18-22 population \approx 47%.

Therefore, an accurate interpretation is that half rather than 90% of the 18-22 years old population were admitted to the universities and colleges. A correct choice of denominator is essential for righteous judgment. This is a typical unwarranted assertion with wrong data support. The misunderstanding is obviously an instance of base-rate neglect (Tversky & Kahneman, 1974, 1982). It is easy to accommodate the discrepancy of opinions by providing up-to-date information with a consensus on the appropriate choice of denominator.

A National Debate on the Curriculum Reform

This is a more complicated type of unwarranted assertion with slim data support for Cluster 2 problems. A national curriculum scheme has been adopted for the nine-year compulsory education in Taiwan over the past decades. Six years ago, a new curriculum was designed in the hope that a traditional teaching scheme of covering more than 20 subject areas could be merged into seven learning areas to ease the learning burden. Whatever merits it may be, a national unrest was first outburst in the fear of the possible worsening of mathematical ability among primary school students due to a six-year learning under constructivistic instruction.

Most disputes centered on the teaching difficulties following a drastic change from specialized subject areas to a wider coverage of learning areas and a deregulation of the exclusive national standard textbook back to the free-market publications. For those students from poor areas, they can not afford a sufficient number of teachers who are capable to conduct efficient teaching and they are resource-limited to buy as many edited textbook on the market as possible. The arguments thus go like this: The implementation of a new curriculum will enlarge the learning divide among different income group students and, henceforth, a possible worsening of educational achievement against the disadvantaged learners.

How can we confirm the worries so that we can modify the curriculum design to a satisfactory degree? Or, how can we falsify the unwarranted beliefs so that we can lift the unwanted collective anxiety away? The best strategy is to conduct two separate national tests, one on the students with a complete learning of standard curriculum and the other on the students with a reformed curriculum learning. The data collected can be set as a benchmark for critical comparisons. Table 1 lists a proposal for doing comparisons to explore the effect of curriculum reform on the worsening of educational achievement. A comparison between 2003 and 2007 TIMSS ranking results is best suited for the purpose. However, it takes too long to ease the unrest. A compromised action was administered by the National Science Council to conduct an IAEP-matching Mathematics Test in 2003 and 2004 respectively. It showed that no statistically significant differences in Mathematics achievement can be found between two groups of students who are subject to different curriculum learning (Huang, 2004).

Table 1 A Critical Administration of Achievement Tests Among Eighth Graders for the Verification of Curriculum Reform Effects in Taiwan

Tests	Results	Curriculum Types Received at Their Inception of Tests
1999 TIMSS-R* ¹	Taiwan ranked the third on Mathematics and the first on Science	Standard curriculum
2003 TIMSS	Taiwan ranked the fourth on Mathematics and the second on Science	Standard curriculum
2007 TIMSS	In preparation	Reformed curriculum
1991 IAEP* ²	Taiwan ranked the second on both mathematics and science	Standard curriculum
2003 IAEP-matching Mathematics Test* ³	Better than what performed on 1991 IAEP test in both content and cognitive domains on mathematics	Standard curriculum
2004 IAEP-matching Mathematics Test	No statistically significant differences were found between 2004 and 2003 IAEP-matching test	Reformed curriculum in junior high school learning and completed the primary school learning with six-year constructivistically designed mathematics

*¹ The Third International Mathematics and Science Study-Repeat (TIMSS-R), administered by the IEA.

*² The International Assessment of Educational Progress (IAEP), administered by the Education Testing Service.

*³ A matching mathematics test was designed and intended to be compatible with the IAEP test construction in both content and cognitive domains, administered by the National Science Council. Where content domain is composed of number, measurement, geometry, data, and algebra; while cognitive domain comprises using concepts, knowing facts and procedures, and solving routine problems.

Local Equilibrium for Cluster 3 Problems

For those debates like Cluster 3 problems, they all need to do a lot of efforts for consensus seeking. However, to find a resolution to resolve differences under multiple criteria is not at all easy. A tentative discussion will be postponed to the next section. We will be satisfied here to take tracking issue as an example.

Tracking or ability grouping is a practice of grouping students of similar ability or prior achievement together for instruction. Many countries include both tracked and untracked high schools in their educational systems. Most junior high schools in central and southern Taiwan also adopted tracking to group students between classes. Several decades of research and debate have yet to prove that one is better than the other on the criteria of both efficacy and fairness (Loveless, 1998). Dickens & Flynn (2002) claims that if social multiplier effects and gains from an appropriate curriculum were both at work, normal tracking produces little or no decline in achievement test scores for less able children and perhaps some gains for the more able. However, it's usually the case in Taiwan that less able children can not gain from the fixed curriculum due to a joint entrance examination ahead but lose from lack of contact with the more able.

To achieve a local equilibrium, two alternatives can be evaluated for adoption. The first alternative assumed a tracking policy and should concentrate on improving the quality of education in both settings. However, a once-popular practice of tracking is now prohibited by law in Taiwan. Even if tracking were permitted, it is still very difficult to monitor if the equivalent quality of education has been achieved. In most cases, they are not. The other alternative seems to be more plausible as a local equilibrium solution. The school should be detracked in principle to conform to the national regulation. However, a limited-scale of ability grouping can be applied to the learning of English and mathematics. The grouping of students can only be performed and limited to two or three classes instead of a re-grouping of the whole range of graders. At the same time, quality teaching with matching-level study materials must be assured with constant monitoring.

Consensus Seeking Under Multiple Criteria: The Role of Policy Response

The educational debate, especially for those of Cluster 3, can be described as a simple linear system as Richardson (1939) had long suggested. It reads as follows:

$$\frac{dx}{dt} = ky - \alpha x + g, \quad \frac{dy}{dt} = qx - \beta y + h \quad \dots\dots\dots (1)$$

Where x denotes the debate potential of concerned groups in a generic sense, and y denotes that of the administration or, specifically, the Ministry of Education. The rate of change of $x(t)$ depends on $y(t)$ and on the uneasy feeling of concerned groups towards a lack of sincerity in the administration. We represent these terms by ky and g respectively, where k and g are

positive constants. These two terms cause x to increase. On the other hand, the cost of debate has a restraining effect on dx/dt . We represent this term by $-ax$, where a is also positive. A similar analysis holds for dy/dt and h denotes a negative feeling towards the opposition tenets.

The equilibrium solution of $x(t)$ and $y(t)$ in this system is stable if $\alpha\beta > kq$. The system behavior tends to be unstable if $\alpha\beta < kq$.

In most cases, the administration is always trying to cool-down the debate, then $k > \alpha$ and $\beta > q$. Let $k = m\alpha > \alpha > 0$. If $1 < m < \beta/q$, then

$$\alpha\beta - kq = \alpha\beta - m\alpha q = \alpha(\beta - mq) > \alpha > 0. \dots\dots\dots (2)$$

The equilibrium solution of the system is stable if $\alpha\beta - kq > 0$ or $\beta - mq > 0$. Therefore, system stability can always be achieved if $\beta/q > m$. The administration could empirically determine with caution where m probably lies along the positive real line. Through a careful selection of m to minimize β/q , the administration (ironically, a more liberal one) can escape from bitter criticism of cowardedness or loss of ideals and avoiding harsh arguments with the concerned groups. In such a push and pull transaction, the cooling-down efforts from the administration site will help lift off unnecessary emotional tension or exaggerated sense of justice and fairness so that a consensus seeking under multiple criteria can be made possible. A cooling-down practice is not the substantial part of real consensus seeking, however, it is a prerequisite for initiating a rational approach to find the genuine equilibrium.

References

- de Queriain, D.J.-F., et. al. (2004). The neural basis of altruistic punishment. Science, 305, 1254-1258.
- Cave, P. (2001). Educational reform in Japan in the 1990s: “individuality” and other uncertainties. Comparative Education, 37, 173-191.
- Dickens, W.T., & Flynn, J.R. (2002). The IQ paradox is still resolved: Reply to Loehlin (2002) and Rowe and Rogers (2002). Psychological Review, 109, 764-771.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. Nature, 425, 785-791.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C.F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. Science, 310, 1680-1683.
- Huang, J.T. (2004). A National Debate on Curriculum Reform And the Worsening of Educational Achievement. A welcome address presented to the 45th General Assembly of the International Association for the Evaluation of Educational Achievement (IEA), October 11-14, 2004, Taipei, Taiwan.

- Loveless, T. (1998). The tracking and ability grouping debate. Washington, D.C.: Thomas B. Fordham Foundation.
- Pan, H.L., & Yu, C. (1999). Educational reforms with their impacts on school effectiveness and school improvement in Taiwan, R.O.C. School Effectiveness and School Improvement, 10, 72-85.
- Richardson, L.F. (1939). Generalized foreign politics. The British Journal of Psychology, Monograph Supplement #23.
- Tetlock, P.E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. Psychological Review, 109, 451-471.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.
- Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases, Ch. 10. Cambridge: Cambridge University Press.

計畫成果自評

1. 本研究報告內容與原計畫書大體相符，先經由蒐集 595 篇論文及評論文章予以分為三大類，並從行為決策角度作詳盡分析，與原計畫書所提各項皆稱符合，達成預期目標。
2. 本研究成果以行為決策理論分析國內以 2002-2004 為主之教育爭議類型，並說明三大類型爭議之政策因應方式，應具有學術與實用價值。
3. 本研究所作之各項分析與結論，甚少出現在同類教育研究之中，應適合投稿教育類之學術期刊。