

# The contribution of Jean Cardinet to the field of educational measurement

*Dany Laveault  
Faculty of Education,  
University of Ottawa*

Chers collègues,  
Dear colleagues,  
Cher Jean,

It is certainly a great pleasure and an honor to address you on the contribution of Jean Cardinet to the field of educational measurement. Needless to say, I have not been able to cover the whole “universe” or population of Jean Cardinet’s work. I am not even sure that the universe of Jean Cardinet’s contributions is finite or infinite. Sampling was my only way to get a descent job done. But, then which sampling method should I use?

For such an important undertaking, I could not trust random sampling. Instead, I chose a purposeful sampling method which Canadians have a lot of experience with: that is “snowball sampling”. Its principle is easy: you start with a few preliminary sources, either publications or persons with the information you need, and as one reference leads you to another and one person to yet another one, you roll your ball of information until it gathers enough material to make a meaningful sample or “snowman”. This way I was able to identify important papers and aspects of his work. I read not only from Jean Cardinet but on Jean Cardinet, thanks to the assistance of some of his colleagues. This way, I have been able to gather a collection of information which I intend to articulate in the best way I can to underline some of his contributions deemed appropriate for this conference.

I will organize my presentation along the two roles Jean Cardinet assigned to educational measurement: to inform and to support. Thus, his work encompasses two preoccupations: the requirements for reliable and valid data to inform and the appropriate use of such information to support decision making at all levels of the educational system. His special attention to the “consequential validity” of educational testing probably results from his own realization, as an industrial psychologist at the start of his career, of the consequences of inadequate decisions based on tests results. I will try to articulate his concerns and show how his work provides us with some insights into today’s challenges

## ***The information function of educational measurement***

The expansion of educational measurement is a constant feature of the work of Jean Cardinet. It has been expressed in his 1977 paper presentation at a meeting of educational researchers from Belgium and Switzerland. In the conceptual framework he developed, the generalizability of educational measurement occupies a position of choice. Traditionally, G-studies were used “to anticipate the full variety of designs that must be

used for a D-study”. As Crocker & Algina (1986) have reiterated, “it is not possible to classify a G-Study or a D-study based on its design alone; the purpose of the investigator is the determining factor” (page 158). A major contribution of Cardinet’s work has been to increase the number of situations where G-studies could serve to inform our decisions. With the help of Yvan Tourneur, Linda Allal and several other colleagues, Jean Cardinet (1985) has shown that G-studies could be applied to differentiate not only among subjects, as in Cronbach’s original inception of G-Study, but also among facets such as items, learning objectives and so on. In his 1994 paper entitled “Control of the value of an intra-subject measurement design”, he showed the significance of generalizability theory to measure student progress. His emphasis on the need to expand the domain of educational measurement is clear in this quote:

*The concept of facet slicing ... expands even further the types of applications and optimization that are made possible using generalizability theory. New measurement designs are offered, at the same time intra-individual and based on large scale studies, promising educational benefits completely different from those of traditional psychometrics, while belonging to the same conceptual framework.*

His contribution, as outlined in the previous quote, has consisted in developing new applications for generalizability studies. Reciprocally, such applications have contributed to differentiate and expand the possibilities of testing and to set appropriate reliability requirements for different applications. Some of his conclusions were sometime met with dead-ends. In some instances, the requirements were found to be practically impossible to meet. So many items would be needed to meet descent reliability requirements, that educational measurement would become unworkable. “Some time must be left out for student study periods” said he. When such events occur, as in student assessment, Jean Cardinet “temporarily” leaves the world of generalizability theory to enter the world of qualitative assessment to support learning, also known as “formative assessment”.

Jean Cardinet’s work attests to the importance of both qualitative and quantitative methods. “The real is qualitative” says Jean Cardinet in a paper where he tries to delimit the respective domains of both quantitative and qualitative research methodologies. The quantitative research methodology advantage, on the other hand, lies in its capacity to measure the uncertainty of our decisions. Jean Cardinet’s work borrows from both research epistemological traditions. In some way, he referred to “mixed research methodology” before the term and concept were actually coined. As he showed, quantitative research methodologies limit our affirmations and question our own sense of evidence. That is why generalizability theory is never too far away in several of Jean Cardinet’s papers.

An expansion of the field of educational measurement as exemplified in Jean Cardinet’s work is still much needed today. Recent developments in the field of educational measurement are putting our test models to the test. Here is a short description of these developments and the challenges they present us with:

1. Case # 1. The requirements of authentic forms of assessment and of performance-based testing have raised important questions on the most valid and reliable methods

to code professional judgments while retaining their value of authenticity. What tools and what methods lead to the most generalizable expert conclusions? How many observations by how many judges are necessary to generalize the results of such forms of assessment?

2. Case # 2. The assessment of complex objects such as competencies, either through traditional testing methods or performance-based assessment, presents educational measurement with still other challenges. In a literacy tests, items referring to the same text are not independent and are more appropriately described as “testlets” (Wainer). The choice between analytical or global rating scales is not clear. Integration of different subject matters also brings about special difficulties. All such cases call for appropriate measurement designs and for information on the generalizability of results under such conditions.
3. Case # 3. There is also an increased interest in standard setting methodologies and on their reliability and validity. Reckase (2006) recently compared a modified Angoff method with the booklet method using error models. What is the generalizability of standards produced by different standard setting methods? How efficient are they in terms of number of observations and number of judges required to reach standards we can trust?
4. Case # 4. There is an awareness of the need to integrate cognitive models into educational measurement models. Several statistical latent-trait models have been developed to that effect. Most of them rely on expert judgments to build a Q matrix that connects individual items to a cognitive attribute. So far, the review of literature on cognitive modeling does not offer much direction as to how many experts are needed to build a generalizable Q matrix. If Q matrices are not generalizable, then the chances are that cognitive attributes cannot be linked correctly to a set of items, leading to different diagnoses and prognoses from different samples of experts.

All those recent developments involve an input of experts’ professional judgments at one point or another. We are far from the requirements of selected response assessment and the so-called “objective testing”. In all four previous cases, we need to know how many experts are required under certain conditions to make “generalizable” judgments. Although such issues are not the exclusive domain of generalizability theory, they certainly define a set of measurement problems where G-studies can help to make appropriate decisions.

The important challenges facing educational measurement today echo concerns which were articulated in Jean Cardinet’s 1987 paper entitled “External, internal, or negotiated assessment?”. He then shows the limitations of both the experimental and the social science models which have shaped the role and use of educational measurement. Although he does not discard them altogether, he demonstrates how both models must be complemented by a third model, which he describes as “negotiated assessment”. To quote Jean Cardinet (page 9):

*... assessment necessarily has multiple points of reference, since it takes the combined reference points of the individual protagonists as its framework. Reality is thus defined as the sum-total of objective data from the first approach and the subjective perceptions from the second. The best way of exploring this reality is to*

*compare as many view points as possible and, for this purpose, to generate interaction between the protagonists.*

The value of assessment is not to be found solely in its capacity to control the educational system, as in experimental sciences, nor solely in its capacity to understand and explain it, as in social sciences, but ultimately in its capacity to produce changes and bring in innovation to improve the school system. As I said initially, “consequential validity” is a core concern in Jean Cardinet’s work. Such a concern is not limited to the information function of educational measurement. It is also present in its function of support. The constructive application of measurement to educational issues involves the human factor to a greater extent. That is where the work of Jean Cardinet leaves the secure world of generalizability theory to enter the most unsteady ground of human decision making.

### ***The support function of educational measurement***

Statistics have been labeled “the most inexact science of the exact sciences”. Some people go as far as to pretend that God created the statistician to make the weatherman look good. When contemplating useful applications of educational measurement, we necessarily reach a point where we must move beyond our measurement models. Such a move toward the integration of external models is necessary to improve our actions whether such actions have to do with measurement itself, its communication or its application. Another contribution of Jean Cardinet’s work has been to enrich the field of educational measurement with contributions from different disciplines such as cybernetic, theories of system, piagetian theory, social psychology, epistemology and ethics to name a few. The expansion of educational measurement cannot be accomplished without a corresponding expansion of its theoretical base.

Jean Cardinet’s work sets high standards for test developers and test users. In his 1977/1986 paper “Educational objectives and individualized assessment”(Objectifs éducatifs et évaluation individualisée) , he assimilates test development and the use of test results to “rocket science”. Using a cybernetic model, he draws a parallel between a rocket launch to planet Mars and the different feedback loops and regulations occurring at different points in a system of educational evaluation. Educational assessment involves a series of complex operations which must occur in a timely manner for the mission to succeed. Its mission has three main targets: improve conditions of learning, improve learning processes and finally improve learning outcomes. Of particular concern is the confusion of perspectives which often results in the use of inappropriate “control” of testing instruments or interpretation of results. Using a major concept from cybernetic, the concept of regulation, he applies it to issue of individualized assessment, emphasizing the primacy of its formative function. Since his 1977 seminal paper, the concept of regulation has been expanded by Linda Allal and several other francophone researchers to cover other functions or “missions” of educational measurement.

Thirty years after Jean Cardinet’s Mars mission, what can we say of the educational measurement mission? It seems to me that the Mars mission analogy would fit better within the experimental or the social sciences model. To remain coherent with the model

of “negotiated assessment” which Jean Cardinet later developed, I have looked for an analogy which would carry the “negotiated” component occurring in human social interactions. I found such an analogy in the recent climate change negotiations of the Bali conference. On several accounts, the whole issue of climate change and of the necessary course of actions that should be undertaken to limit its negative impact, reminds us of the course of actions undertaken by the collective of teachers and school administrators to change school climate to effect student learning. Here are some of the conditions occurring in both “climate change negotiations”

- People may not agree on the complex nature of the problem and sometime, even question its existence.
- Even among people who agree on the existence of a problem, there may be disagreement on what the targets should be and what are the best means to reach them. Some participants will rightfully point out the deficiencies in the science upon which decisions are based. Such arguments will often be used to justify that no actions are taken or be rushed to.
- Even among those who agree on the actions to be undertaken, some will not go ahead unless they are pressured to do so. Some people will voluntarily take action, some others will not. Matters of leadership will become central.

A lot of “subjectivity” and personal interest are involved in such negotiations. Instead of ignoring subjectivity or social interactions, Jean Cardinet suggests that we take them into account and use them for the best. He uses the subjectivity inherent in decision making in a way that reminds me of the aikido Japanese martial art. Aikido consists in using the strength of one’s opponent to make him lose his balance and control him in a way he no longer is a threat. In Cardinet’s “negotiated assessment”, an interactionist model of school management, it is hoped that the shared “subjectivities” of individuals will bring about collective representations and values which would have an almost objective status.

Jean Cardinet’s model of negotiated assessment anticipated on some of the most recent concepts and ideas on educational change. For instance, Michael Fullan’s model of “community of learning” is much the same as Cardinet’s model of “negotiated assessment”. Here is a summary of the key features of Fullan’s model (2001) as summarized by Bruce Hammonds (2002):

*The key to successful change is the improvement in relationships between all involved and not simply the imposition of top down reform. The new emphasis in educational change is based on creating the conditions to develop the 'capacity' of both organizations' and individuals to learn. The focus moves away from an emphasis on structural change towards changing the culture of classrooms and schools, an emphasis on relationships and values.*

*... The challenge is how to share and sustain ideas about change so as to transform what is essentially a conservative system Teachers and schools need to be seen as 'moral change agents' making democratic communities possible ..*

<http://www.leading-learning.co.nz/newsletters/vol01-no03-2002.html>

## ***Conclusion : Where are we today?***

The conference title asks : “Where are we today?”. In 1977, Jean Cardinet called for an expansion of the field of educational measurement. At the time, he outlined new directions of potential interests for the field and pointed out some dangerous landmines. Undoubtedly, the field of educational measurement has made tremendous progress. The presentations of this conference on standard setting methods, generalizability theory and item response models attest to such a progress. The field has grown up in assurance and is tackling problems and taking up challenges one could only dream about years ago.

That being said, educational measurement today may also be victim of its own success. Today is a time of high expectations toward educational measurement and of high consequences. The attention it is getting may be disproportionate to its real significance. In his 1977 paper on the coordination of information in the educational system (*La coordination de l'information dans le système éducatif*), Jean Cardinet remarks that the valid use of educational measurement to improve our educational systems is just as essential as the collection of reliable and valid measurement data. High expectations and high consequences call for high standards in the use of such results. Educational measurement specialists have the responsibility to inform test users of the limitations of their instruments and of the conclusions that can be drawn from them. Assuming we could develop perfectly valid and reliable test results, one cannot guarantee that valid and fair decisions would be made unless high standards are being followed, not only by tests developers but by test users as well. The following Ontario and U.S. cases are just a few examples of the bad consequences of not doing so.

- Case #1. The introduction of a provincial testing in Ontario Grade 3 and Grade 6 has sparked unwarranted fears among teachers that student results on such tests could be used to hold them solely accountable. Experienced teachers of grade 3 and grade 6 who had first-choice in the selection of the grades they teach each year, fled to teach other grade levels. After a few years, Grade 3 and Grade 6 students were taught by significantly less experienced teachers as compared to other grades. As a result, student achievement stagnated and even decreased.
- Case #2. Although the administrative process and financial resources necessary to put in place a national program of testing did not take much time to implement once Ontario legislation was voted, it took several years before teachers, school principals and School Boards administrators start paying attention to test results and used them properly. The communication of results in a way that would support administrative decisions and stimulate professional exchanges still present issues which are just as equally important as the tests' results reliability and validity.
- Case #3. Despite all disclaimers, unauthorized comparisons and decisions are still being made by professionals and by non professionals. Ontario newspapers compare results of tests administered in French to results of tests administered in English even if the tests' results are not metrically equivalent. Stakeholders will invariably pressure for different applications of test results, applications which are not always warranted by the test reliability or validity for such purposes.

- Case # 4. In neighboring United States, school ranking resulting from the use of national and state testing has had severe impact not only on educational decisions, but also on the household market. Real estate agents currently use school rankings to justify higher home property values.

Considering the new challenges facing educational measurement, today's conference theme is timely. I wish it will provide you with the same enriching and learning experience I had while reviewing Jean Cardinet's work. I have done my best to select what appeared to be the most meaningful pieces of work from his work. Thanks to the help of some of his colleagues, my decisions have been much easier to make. May they receive here the full extent of my gratitude.

As a final conclusion, I could not end my presentation without saying a few words on Jean Cardinet, the man of conviction and the man of action. His work, as I have tried to illustrate, sets high standards for our discipline. Though his own achievements, he has contributed to expand the field of educational measurement in several ways. You have heard today about some of his ideas. I should end by saying a few words on his actions.

Jean Cardinet did not only set high standards for educational measurement, he also set high standards for himself. He must be recognized for his contribution to the expansion of educational measurement while at the same time maintaining a coherent perspective of the field. One of his colleague gently whispered to my ear a famous quote from Socrates : "Nothing is more socially subversive than someone who compels us to coherence". *Rien n'est plus subversif dans une société que quelqu'un qui oblige à la cohérence* : As a man of coherence and of conviction, Jean Cardinet stands up to the numerous challenges presented by educational measurement and has translated his convictions into actions.

On several accounts, his actions have made a difference for the collective good. Let me mention just a few:

- Jean Cardinet has contributed to the foundation of the International Test Commission, an organization committed to develop international standards on testing. In a paper published in the first issue of the International Journal of Testing in 2001, Oakland, Poortinga, Schlegel and Hambleton recognize the important contribution of Jean Cardinet to the conception of this international organization. It all started with Jean Cardinet's 1968 article on the applications of ethical standards in testing to the General Assembly of the Swiss Psychological Society. Such standards are essential to the permanent progress of our discipline.
- Jean Cardinet has also played a decisive role in the extension of the belgo-swiss meetings and in the development of an international association of francophone experts in the methodologies of evaluation, ADMÉE-Europe, which celebrated its 21<sup>st</sup> year of existence this year. He also played an active role in la *Revue de mesure et évaluation en éducation*, one of the rare scientific French publication of international status in the field of education.

For all that, congratulations to you Jean and Merci!

## References

- Cardinet, J. (1994). Control of the value of an intra-subject measurement design. In D. Laveault, D., B. Zumbo, M. Gessaroli, & M. Boss (Éds). (1994). Modern Theories of Measurement: Problems and Issues. Ottawa: Faculty of Education, University of Ottawa.
- Cardinet, J. (1990/10987). Évaluation externe, interne, ou négociée? Neuchâtel : IRDP, 1987, 17 p. Repris dans Hommage à Jean Cardinet, Cousset (FR) : Delval, 1990, p. 139-157.
- Cardinet, J. (1986/1977). Objectifs éducatifs et évaluation individualisée. Neuchâtel : IRDP, 1977, 57 p. Repris dans «Évaluation scolaire et Mesure», 1986
- Cardinet, J. (1986) Évaluation scolaire et pratique. Bruxelles : De Boeck.
- Cardinet, J. (1986). Pour apprécier le travail des élèves. Bruxelles : De Boeck.
- Cardinet, J. (1986). Évaluation scolaire et mesure. Bruxelles : De Boeck.
- Cardinet, J. & Tourneur, Y. (1985). Assurer la mesure. Berne : Peter Lang.
- Cardinet, J. (1977). La coordination de l'information dans le système éducatif. Neuchâtel : IRDP, 1977, 12 p. In : Actes du 4ème Congrès International de l'AIPELF, Genève, 16-20 mai 1977. Genève : SRP, 1978, p. 149-159.
- Crocker, L. & Algina, J. (1986). Introduction to Classical and Modern Test Theory. New York: Holt, Rinehart and Winston.
- Fullan, M. (2001). The new meaning of educational change: Teachers College Press, 297 pages.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). Washington, DC: American Council on Education & National Council on Measurement in Education.
- Oakland, T., Poortinga, Y.H., Shclegel, J. & Hambleton, R.K. (2001). International Test Commission: Its History, Current Status, and Future Directions. International Journal of Testing, 1(1), 3-32.



## Bibliography

*For most recent references of Jean Cardinet :*

[http://www.unige.ch/fapse/SSE/groups/life/descriptif\\_membres/Fiches\\_MA/Cardinet.html](http://www.unige.ch/fapse/SSE/groups/life/descriptif_membres/Fiches_MA/Cardinet.html)

*For references of his work at IRDP (mostly in French with English abstract):*

<http://www.irdp.ch>

*For references on Jean Cardinet's work up to 1990 and major contributions in French:*

Institut romand de recherche et de développement pédagogique (1990). Hommage à Jean Cardinet. Neuchâtel : Delval.

### **Selected references in English**

*Selected references indexed in ERIC*

Cardinet, J. (1988). The generalisability of an individual's progress. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 8(4), 247-255. Retrieved

Cardinet, J., & Allal, L. (1983). Estimation of generalizability parameters. *New Directions for Testing and Measurement*, (18), 17-48.

Cardinet, J., & And Others. (1981). Extension of generalizability theory and its applications in educational measurement. *Journal of Educational Measurement*, 18(4), 183-204.

Cardinet, J., & And Others. (1979). *Equality of opportunity reconsidered: Values in education for tomorrow. educational research in europe*The Swets North America Inc., P.O. Box 517, Berwyn, PA 19312.

Cardinet, J., & And Others. (1976). The symmetry of generalizability theory: Applications to educational measurement. *Journal of Educational Measurement*, 13(2), 119-135.