OCTOBER 2011

# The Canadian Index of Wellbeing

## TECHNICAL PAPER: CANADIAN INDEX OF WELLBEING 1.0

Alex Michalos, Bryan Smale, Ron Labonté, Nazeem Muhajarine, Katherine Scott, Martin Guhn, Anne M. Gadermann, Bruno D. Zumbo, Alexis Morgan, Kelley Moore, Lenore Swystun, Bill Holden, Heather Bernardin, Beth Dunning, Paul Graham, Ann-Sylvia Brooker, Ilene Hyman

CANADIAN Index OF WELLBEING

Measuring what matters


In referring to this report, please use the following citation format:

Michalos, A.C., Smale, B., Labonté, R., Muharjarine, N., Scott, K., Moore, K., Swystun, L., Holden, B., Bernardin, H., Dunning, B., Graham, P., Guhn, M., Gadermann, A.M., Zumbo, B.D., Morgan, A., Brooker, A.-S., & Hyman, I. (2011). *The Canadian Index of Wellbeing.* Technical Report 1.0. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

# Contents

## List of Tables



## List of Figures

## Executive Summary



The aim of this paper is to describe an approach to the construction of a single composite index worthy of the name "Canadian Index of Wellbeing" (CIW) based on a selection of headline indicators. Imagine, if you can, a single tree capable of capturing some of the magnificence of a Canadian forest and you will be able to appreciate the challenge before us. Technically speaking, it is a task of constructing a unidimensional index to reasonably represent a multidimensional construct of human wellbeing. The paper provides some background material reviewing assumptions made and principles agreed upon by the Canadian Research Advisory Group (CRAG), formerly the CIW National Working Group, in general and the authors of this piece of work in particular.

We regard the diverse features of development work on the CIW as part of a System insofar as all the work done has the potential to make a contribution toward the construction of a single composite index. The detailed indicators and indexes presented in this document are based on commissioned reports for the eight domains of the System. The full reports, listed in the References, are as follows:

- ***Community Vitality*** by Katherine Scott
- ***Democratic Engagement*** by Kelley Moore, Lenore Swystun, Bill Holden, Heather Bernardin, Beth Dunning, and Paul Graham
- ***Education*** by Martin Guhn, Anne M. Gadermann, and Bruno D. Zumbo
- ***Environment*** by Alexis Morgan
- ***Healthy Populations*** by Ronald Labonté, Nazeem Muhajarine, Brandace Winquist, and Jacqueline Quail
- ***Leisure and Culture*** by Bryan Smale, Holly Donohoe, Clem Pelot, Agnes Croxford, and Denis Auger
- ***Living Standards*** by Andrew Sharp and Jean-François Arsenault
- ***Time Use*** by Ann-Sylvia Brooker and Ilene Hyman

Each report has a set of indicators providing a broad description in statistical terms of aspects of the named domains that seem to be of primary concern. From the large number of indicators reviewed, using our accepted assumptions and principles, the CRAG identified a set of headline indicators ("headliners", for short) for each domain. Headliners have the function of making certain aspects of the domains easier to see, of drawing attention to a few key, representative trees, as it were, in a fairly dense forest. Since the primary aim of this paper is to introduce the CIW System and our approach to building a composite index, we anticipate changes in the selection of headliners as discussions about the work completed so far lead to revisions in the future.

In order to provide context to our work, in the first two sections of this paper, we briefly sketch the recent history (past 40 years) of attempts to create composite indexes of quality of life. We discuss the difficulties encountered since the fifth century BCE arising because of the fuzzy edges surrounding ordinary ways of talking about health, wellbeing, the quality of life, happiness, and a good life, all things considered. The logo for the CIW includes a subtitle saying "Measuring what matters", which suggests a broad understanding of wellbeing. We assume that "overall wellbeing" is roughly synonymous with "overall quality of life". While the quantity of our lives is notoriously limited to one per person, its quality is as varied as the perspectives or domains from which it is viewed. Viewed from one perspective, a person may be well off, whereas from another perspective, not at all well off. For example, a poor person might have good family relations and spiritual fulfillment while living in a rough neighbourhood with substandard housing. Someone with a good job may suffer from a long and lonely commute every day. In the context of research on wellbeing or quality of life, "measuring what matters" implies measuring our lives from the perspectives that are most important to us.

The development of the CIW has been and probably will remain pragmatic. Practically speaking, that means that we proceed patiently, transparently, and flexibly, testing any ideas presented both against the hard evidence yielded by empirical research and against the common sense of the CRAG and as broad a constituency beyond it as our resources allow.

It was agreed relatively early that most of the phenomena relevant to human wellbeing at the present time could be conceptualized in eight Domains; that is, living standards, healthy populations, community vitality, democratic engagement, leisure and culture, time use, education, and the environment. It also was agreed that any acceptable indicator or index of wellbeing should be a statistical measure that satisfies some familiar *Acceptability Criteria*, including, for example, relevance to the concerns of our main target audiences (i.e., ordinary Canadians, elected officials, unelected administrators, experts), easy to understand, reliable and valid, and politically unbiased. These agreements were supplemented by a list of *Critical Issues* that creators of all indicators and indexes have to address, such as what spatial and temporal coordinates should be used (i.e., cities, provinces, the country as a whole, over one or more years, beginning and ending when); definitions of domains of concern and diverse, identifiable groups; ways to assess benefits and costs; and validating or auditing criteria for work completed.

There is a substantial body of literature devoted to the question of whether, all things considered, an array or profile of diversely related distinct indicators of wellbeing in separate silos is more useful than a single composite index. We unhesitatingly assert that individual indicators and composite indexes each have advantages and disadvantages that vary with the particular uses to which they are put, as well as with the way they are put, in what circumstances, at what time, with what resources and constraints, and by whom. This mentions only a few of the most obvious conditions for successful or unsuccessful applications. In order to make the set of procedures used to produce and validate the whole System as transparent as possible, its components are listed: commissioned expert literature reviews, validation in the

form of peer reviews and CRAG assessments, the selection of headliners, and an aggregation function.

The technical problem of constructing a unidimensional scale to reasonably represent a multidimensional construct of human wellbeing is solved by creating a *mean of percentage change rate ratios scale* (*percentage change scale,* for short). Because percentage change scales allow trade-offs between deteriorations on some indicators to be compensated by improvements in others, they may be regarded as *compensatory* scales.

Because most of our health statistics were drawn from the various cycles of the National Population Health Survey, which began in 1994, 1994 was selected as our base year. Selection of 2008 as our final year of review was determined by the latest full set of data available across all eight domains. To create comparable index values from our raw data values, the baseline values of each of the 64 headliners are set at 100. Positive percentage changes for each one indicate some improvement or positive change in wellbeing while negative percentage changes indicate some deterioration or negative change.

In considering the assignment of weights in a principled way to each headline indicator, we followed a variant of Laplace's *Principle of Nonsufficient Reason*, which suggests in the absence of a sufficient reason to regard any particular indicator as more important than any other, each indicator should be assigned an equal weight. There are many reasons for regarding one or another indicator as more important in some way or other, but what is missing is a good reason for assigning any particular indicator a particular numerical value greater or less than that of some or all other indicators. The absence of such a reason justifies the equal treatment of all indicators here. With the greater understanding of the relationships among all indicators that is bound to come as research on the CIW proceeds, sufficient reasons for diverse weights may appear.

The aggregation function used for the index values for the eight indicators within each domain as well as for the composite index of the eight domains is a simple average or mean score. The simple average of any set of numbers is a familiar measure of the central tendency of the set, with familiar problems. Most notably, a mean (or average) score can provide a misleading picture if one or a few figures in the set are wildly different from most others.

Table 1 gives an overview of the final average scores for each of the eight domains, the 64 indicators, and the CIW itself. From this table, one can see that there were 39 headline indicators showing improvements or positive changes over the period and 25 showing deterioration or negative change. Over the whole 15 year period from 1994 to 2008, Canadians enjoyed an 11.0% improvement in their overall wellbeing.

The total value of percentage increases of the five indicators with the greatest improvements was considerably higher than that of the five with the greatest deterioration (i.e., 477.7% compared to 157.3%). On the positive side, there was a 160.4% decrease in the percentage of the labour force with long-term unemployment, 106.6% increase in the percentage reporting

that policies of the federal government had made them better off, 83.3% decrease in the percentage of daily or occasional smokers among teens aged 12 to 19 years, 66.7% increase in the ratio of childcare spaces to children aged 0 to 5 years, and a 60.7% increase in the Viable Non-Renewable Energy Reserves Index. On the negative side, there was a 49.2% increase in the percentage of people reporting that they had diabetes, a 37.3% decrease in the Viable Metal Reserves Index, a 25.6% decrease in net Official Development Aid as a percentage of gross national income, a 23.8% decrease in the Canadian Living Planet Index, and a 21.4% decrease in average visitation per site in the past year to all national parks and national historic sites.



**Table 1**
**CIW List of Indicators for All Domains, with Percentage Gains and Losses (1994 to 2008)**

| ***Domain*** <br> Indicator | **Pct. change[a]** |
|---|---|
| ***Community Vitality*** | |
| Percentage reporting participation in organized activities | 27.3 |
| Percentage with 6 or more close friends | 10.1 |
| Property crime rate per 100,000 population | 34.0 |
| Violent crime rate per 100,000 population | 1.1 |
| Percentage who feel safe walking alone after dark | 10.3 |
| Percentage disagreeing that they worry less about the needs of others | 55.6 |
| Percentage who provide unpaid help to others on their own | 15.1 |
| Percentage reporting very or somewhat strong sense of belonging to community | 12.3 |
| **Domain Average** | **20.7** |
| ***Democratic Engagement*** | |
| Percentage of voter turnout at federal elections | -11.8 |
| Percentage that are not interested in politics at all | 36.6 |
| Percentage strongly agree it is every citizen's duty to vote in federal Elections | 14.7 |
| Pct. reporting they are very/fairly satisfied with the way democracy works in Canada | 3.1 |
| Pct. reporting that policies of the federal government have made them better off | 106.6 |
| Ratio of registered to eligible voters | 6.7 |
| Percentage of women in Parliament | 24.4 |
| Net official development aid as a percentage of gross national income | -25.6 |
| **Domain Average** | **19.3** |
| ***Education*** | |
| Ratio of childcare spaces to children aged 0 to 5 years of age | 66.7 |
| Percentage of children doing well on five developmental domains | 3.6 |
| Ratio of students to educators in public schools | 8.2 |
| Average of 5 social and emotional competence scores for 12 to 13 year olds | -3.7 |
| Basic knowledge and skills index for 13 to 15 year olds | -0.2 |
| Percentage of PISA scores explained by socio-economic background | 22.2 |
| Percentage of 20 to 24 year olds in population completing high school | 5.8 |
| Percentage of 25 to 64 year olds in population with a university degree | 47.4 |
| **Domain Average** | **18.7** |

| *Domain*<br>Indicator | Pct. change[a] |
|---|---|
| ***Environment*** | |
| Ground level ozone (population weighted in parts per billion) | -3.5 |
| Absolute GHG emissions (megatons of $CO_2$ per year) | -15.0 |
| Primary energy production (petajoules) | 17.7 |
| Water yield in Southern Canada ($km^3$) | 3.9 |
| Viable Non-Renewable Energy Reserves Index | 60.7 |
| Viable Metal Reserves Index | -37.3 |
| Canadian Living Planet Index | -23.8 |
| Marine Trophic Index | -5.3 |
| **Domain Average** | **-0.3** |
| ***Healthy Populations*** | |
| Percentage self-rated health as excellent or very good | -6.7 |
| Percentage with self-reported diabetes | -49.2 |
| Life expectancy at birth, years | 3.3 |
| Percentage of daily or occasional smokers among teens aged 12 to 19 years | 83.3 |
| Percentage with probable depression | -11.7 |
| Percentage rating patient health services as excellent or good | 2.8 |
| Percentage aged 65 years or more getting influenza immunization | 34.2 |
| Avg. number of remaining years expected to be lived in good health (avg. HALE 15+) | -3.9 |
| **Domain Average** | **6.6** |
| ***Leisure and Culture*** | |
| Average percentage of time spent on the previous day in social leisure activities | -18.5 |
| Average percentage of time spent on the previous day in arts and culture activities | -5.4 |
| Average number of hours in the past year volunteering for culture and recreation organisations | -19.5 |
| Avg. monthly frequency of participation in physical activity lasting over 15 minutes | 24.5 |
| Average attendance per performance in past year at all performing arts performances | 0.5 |
| Average visitation per site in past year to all National Parks and National Historic Sites | -21.4 |
| Average number of nights away per trip in the past year on vacation trips to destinations over 80 km from home | 11.3 |
| Expenditures in past year on all aspects of culture and recreation as a percentage of total household expenditures | 4.2 |
| **Domain Average** | **-3.0** |
| ***Living Standards*** | |
| Ratio of top to bottom quintile of economic families, after tax | -13.9 |
| After tax median income of economic families (2008$) | 24.0 |
| Percentage of persons in low income | 48.9 |
| Scaled value of CSLS economic security | -8.7 |
| Percentage labour force with long-term unemployment | 160.4 |
| Percentage of labour force employed | 8.8 |
| CIBC index of employment quality (1994 QI=100) | -0.8 |
| RBC housing affordability index | -7.7 |
| **Domain Average** | **26.4** |

| *Domain*<br>Indicator | Pct. change[a] |
|---|---|
| ***Time Use*** | |
| Percentage of 20 to 64 year olds working over 50 hours per week | 21.5 |
| Percentage of 20 to 64 year olds reporting high levels of time pressure | -16.3 |
| Percentage of 20 to 64 years old giving unpaid care to seniors | -10.8 |
| Percentage of 65 years and older reporting daily active leisure activities | -4.5 |
| Percentage of 65 years and older reporting annual formal volunteering activities | 9.2 |
| Pct. of 12 to 17 year olds spending two hours or more per day on TV or video games | -14.2 |
| Percentage of 6 to 9 year olds having weekly or more structured activities | 9.0 |
| Percentage of 3 to 5 year olds read to daily by parents | 1.5 |
| **Domain Average** | **-0.6** |
| **Composite Canadian Index of Wellbeing** | **11.0** |

[a] Percentage gains and losses from 100 in base year (1994). Positive values indicate improvement and negative values indicate deterioration.

Figure 1 illustrates the index trends for the period 1994 to 2008 for each of the eight domains, the composite CIW based on those domains, as well as the Canadian GDP per capita. The figure is constructed so that numbers greater than 100 indicate improvements in wellbeing (green lines) and numbers less than 100 indicate deterioration in wellbeing (red lines). GDP per capita increased substantially more than the composite CIW over the 15 year period.

The living standards domain improved relatively more than all others. The trends in the four domains below the composite CIW throughout most of the period show that deterioration in leisure and culture, time use, the environment, and healthy populations tended to decrease the CIW while improvements in living standards, community vitality, democratic engagement, and education tended to increase it. Assuming all domains are equally important to our wellbeing, we seem to have as much reason to focus on improving the domains where we are relatively strong or on those where we are relatively weak.

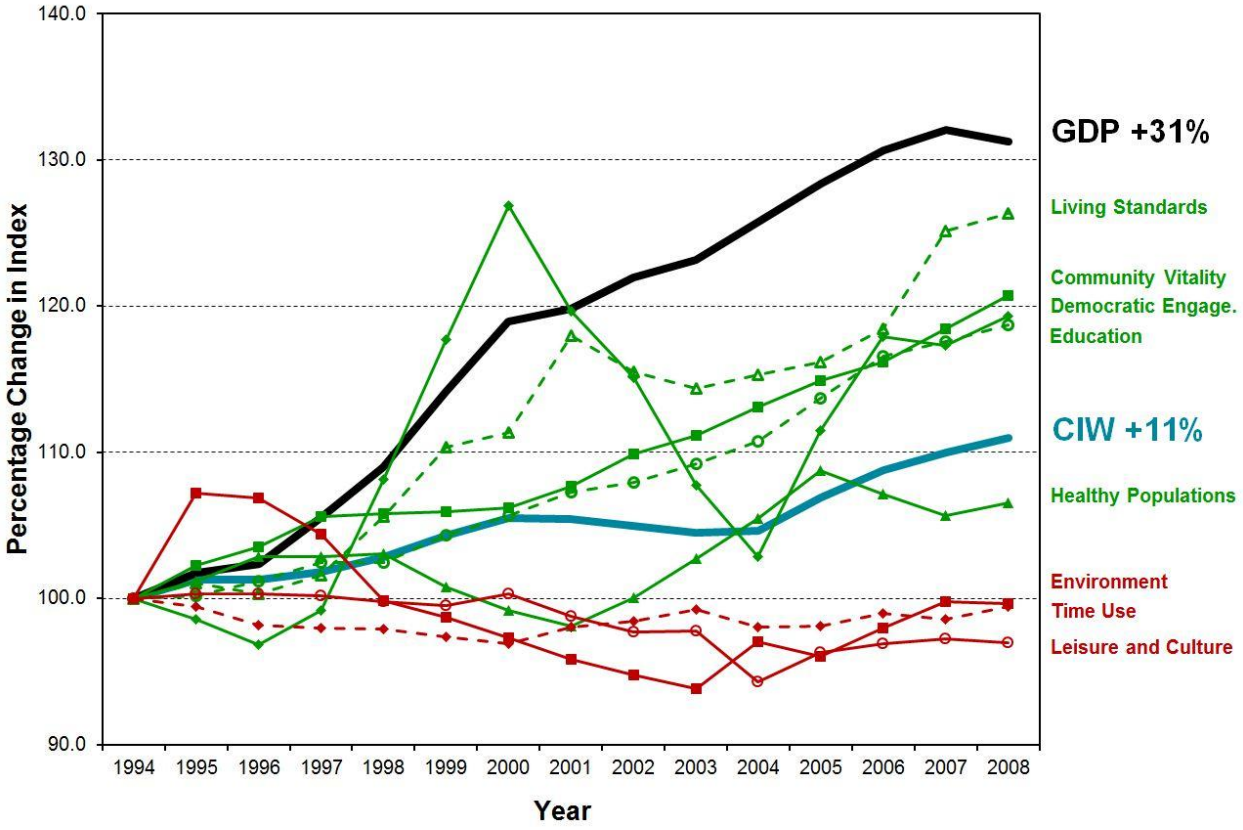
We hope that this paper will encourage others to take up the challenge of creating the sort of comprehensive system and composite index envisioned here. The absence of adequately resourced specific programs of development for a comprehensive system and composite summary index reduces the chances of researchers to raise a number of important questions including:

- the completeness or incompleteness of current stocks of statistical time series,
- links in the form of causal interactions or mere correlations among the indicators housed in different silos,
- the collection of redundant indicators needlessly absorbing scarce resources,

- the failure to collect important data whose availability might reveal serious limitations and/or distortions of our understanding of the quality of our lives, and
- the consequences of making public policy on the basis of research relying on weak conceptual frameworks and/or incomplete data.



**Figure 1**
**Trends in the Canadian Index of Wellbeing and its Eight Domains, Compared with GDP (1994 to 2008)**



# Introduction

In the 1960s, there was a flurry of action among researchers around the world who were fed up with what Bertram Gross (1967) called "the New Philistinism" that seemed to dominate policy makers' attention at the time. The most frequently used measures of the progress of societies and the quality of life of people were financial, with the National Income and Product Accounts, and more particularly, its flagship composite index, Gross Domestic Product (GDP), bearing most of the burden. The diversity of dimensions "summarized" in GDP covers "special – and often heroic – assumptions" (Sen, 1999, pp. 76-119) concerning what McMurtry (1998) called the "life-blind" metric of exchange value and its use as a measure of quality of life.

Those joining the effort to develop measures that were sensitive to many more things that matter became part of what was called "the social indicators movement". Social indicators are statistics that are supposed to have some significance for measuring the quality of life or human wellbeing broadly construed. The movement grew along familiar lines of development, beginning with a few likeminded pioneers who wrote and exchanged papers, organized symposia, established a scholarly journal (*Social Indicators Research*), and later professional societies. By the end of the 1980s, some scholars pronounced the movement dead, since official government agencies like Statistics Canada and international agencies like the Organization for Economic Cooperation and Development (OECD) withdrew resources that had earlier generated so much activity (Land, Michalos, & Sirgy 2011; Sirgy, Michalos, Ferriss, Easterlin, Patrick, & Pavot 2006).

As we entered the twenty-first century, the movement seemed to be resuscitated. Two new professional organizations were formed – the International Society for Quality of Life Studies and the International Society for Quality of Life Research. The OECD launched *its Global Project on Measuring the Progress of Societies* that led to the World Forum on Statistics, Knowledge and Policy in Istanbul, which subsequently led to the *Istanbul Declaration* of June 30, 2007. The declaration began as follows:

> We, the representatives of the European Commission, the Organization for Economic Cooperation and Development, the Organization of the Islamic Conference, the United Nations, the United Nations Development Programme and the World Bank, recognize that while our societies have become more complex, they are more closely linked than ever. … We are encouraged that initiatives to measure societal progress through statistical indicators have been launched in several countries and on all continents. Although these initiatives are based on different methodologies, cultural and intellectual paradigms, and degrees of involvement of key stakeholders, they reveal an emerging consensus on the need to undertake the measurement of societal progress in every country, going beyond conventional economic measures such as GDP per capita. (OECD, 2007)

In 2008, Nicolas Sarkozy, President of the French Republic, established the Commission on the Measurement of Economic Performance and Social Progress (Commission, 2009). It was headed

by two Nobel laureates, Joseph Stiglitz and Amartya Sen, and coordinated by the French economist, Jean-Paul Fitoussi. The 22 members of the Commission included three other Nobel laureates in economics, a psychologist, a political scientist, and 15 other economists. The first third of the Commission's report reviewed the strengths and weaknesses of GDP, the second reviewed some of the work on quality of life measurement over the last 30 years, and the third highlighted some problems of measuring sustainability. The stature of the sponsors and the economists who produced the report, the content of the report itself, and the importance of the organizations that endorsed the *Istanbul Declaration* signify that the economic paradigm of progress is at least severely wounded if not dead. The gigantic task before us is to create statistics that adequately, reliably, and validly measure progress in a comprehensive way, a way revealing the wellbeing or quality of our lives. It is, after all, life of a good quality that we all want to create and ultimately sustain.

The aim of this paper is to describe an approach to the construction of a single composite index worthy of the name "Canadian Index of Wellbeing (CIW)" based on a selection of headline indicators. It is our attempt to create the comprehensive measuring instrument that was suggested in the preceding paragraphs. Imagine, if you can, a single tree capable of capturing some of the magnificence of a Canadian forest and you will be able to appreciate the challenge before us. Technically speaking, it is a task of constructing a unidimensional scale to reasonably represent a multidimensional construct of human wellbeing. The paper provides some background material reviewing assumptions made and principles agreed upon by the CIW National Working Group in general and the authors of this piece of work in particular. To keep the paper to a manageable length, we have omitted all the philosophical and methodological discussions leading to our final selection of the assumptions and principles recorded here.

We regard the diverse features of development work on the CIW as part of a System insofar as all the work done has the potential to make a contribution toward the construction of a single composite index. The detailed indicators and indexes presented in this document are based on commissioned reports for the eight domains of the System that are listed in the References. On average, the reports run about 100 pages each. Briefly, they are as follows:

- ***Community Vitality*** by Katherine Scott
- ***Democratic Engagement*** by Kelley Moore, Lenore Swystun, Bill Holden, Heather Bernardin, Beth Dunning, and Paul Graham
- ***Education*** by Martin Guhn, Anne M. Gadermann, and Bruno D. Zumbo
- ***Environment*** by Alexis Morgan
- ***Healthy Populations*** by Ronald Labonté, Nazeem Muhajarine, Brandace Winquist, and Jacqueline Quail
- ***Leisure and Culture*** by Bryan Smale, Holly Donohoe, Clem Pelot, Agnes Croxford, and Denis Auger
- ***Living Standards*** by Andrew Sharp and Jean-François Arsenault

- ***Time Use*** by Ann-Sylvia Brooker and Ilene Hyman

Each report has a set of indicators providing a broad description in statistical terms of aspects of the named domains that seem to be of primary concern; that is, the indicators represent our efforts to measure what matters most in each domain. From the large number of indicators reviewed, using our accepted assumptions and principles, the Working Group identified a set of headline indicators ("headliners", for short) for each domain. Headliners have the function of making certain aspects of the domains easier to see, of drawing attention to a few key, representative trees, as it were, in a fairly dense forest. The inclusion of all the headliners proposed by the authors of the domain reports in the composite CIW would have made the latter unmanageably large. Thus, some reduction of all the candidate headliners had to be made by the authors of this report, not all of whom agreed with all of the selections finally made. Since the primary aim of this report is to introduce the CIW System and our approach to building a composite index, we anticipate changes as discussions about the work completed so far lead to revisions in the future.

In the next section we provide an overview of definitional problems concerning the central topics of this project (i.e., wellbeing, quality of life, and health). This is followed by a broad characterization of three approaches to the development of indicators and indexes. Next, there is a section giving a list of *Acceptability Criteria* for any indicators and indexes, and then a list of *Critical Issues* that all developers of such indicators and indexes must address. The items in these lists are probably not exhaustive of all those that might be articulated, and in pairs many are certainly not mutually exclusive. Focusing directly on composite indexes, there is a section listing their purported advantages and another listing their purported disadvantages. Following these reviews, we explain some basic concepts of validation and describe the specific validating procedures used for the CIW System. We then give a description of our recommended approach to a composite index of wellbeing, some illustrative headliner data, comment on trends revealed by the composite, and remark on some problems still lacking entirely satisfactory solutions.

## Wellbeing, Quality of Life, and Health

At least since the fifth century BCE, the Greek physicians and philosophers who wrote about human wellbeing, the good life, and health freely used a vocabulary that is similar to that we use today, with the same fuzzy edges surrounding most of the key words (Michalos & Robinson, 2011). Aristotle was the most famous philosopher to come out of that period and in his classic *Nicomachean Ethics*, he wrote:

> What is the highest of all the goods achievable in action? As far as the name goes, most people virtually agree: for both the many and the cultivated call it happiness, and they suppose that living well and doing well are the same as being happy. But they disagree about what happiness is, and the many do not give the same answer as the wise. (Aristotle, 1999, p.3)

In fact, as Michalos and Robinson (2011) showed, many of the wise of that period did not give the same answer as others equally wise.

The disagreements today are as profound as they were 2500 years ago. When the World Health Organization (WHO) defined "health" as "complete physical, mental and social well-being" in 1946, they did not bother to define "wellbeing". In the *Ottawa Charter for Health Promotion* in 1986, the WHO claimed that "Good health is a major resource for social, economic and personal development and an important dimension of quality of life". When contemporary researchers working in the field of health-related quality of life use health status measures to assess the quality of people's lives, they are caught in a viciously circular argument. If good health equals a good life, then the phrase "health-related quality of life" means either "health-related health" or "quality of life-related quality of life" (Michalos 2004; Michalos, Ramsey, Eberts, & Kahlke, 2011). The idea that "good health is … an important dimension of quality of life" is therefore completely undermined, and the otherwise reasonable research question of the relative impact of good health on a good life becomes unreasonable.

A relatively famous quotation from the report of the World Commission on Environment and Development, *Our Common Future* (1987, pp. 43-44), connects sustainable development directly to quality of life.

> Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs ... The essential needs of vast numbers of people in developing countries – for food, clothing, shelter, jobs – are not being met, and beyond their basic needs these people have legitimate aspirations for an improved quality of life.

As this sentence is being written, we have just passed the middle of the United Nations Decade of Education for Sustainable Development (DESD), 2005 to 2014. In the *Framework for the UNDESD International Implementation Scheme*, one finds:

> The overall goal of the DESD is to integrate the values inherent in sustainable development into all aspects of learning to encourage changes in behavior that allow for a more sustainable and just society for all … It is essential to situate the Decade with respect to efforts in which the international community is already engaged … All of [these efforts] aim to achieve comparable impacts: an improvement in the quality of life… (UNESCO Education Sector, 2006a, 2006b, pp. 3-11)

This brief review of concepts concerning wellbeing, quality of life, health, and sustainable development shows that there are significant overlaps among these ideas. In a somewhat heroic effort to bring more precision than that displayed by ordinary language regarding these ideas, 52 contemporary and well-known scholars signed on to a document intended to at least eliminate some of the apparent fuzziness. Agreement was possible primarily because most of those scholars recognized the problems created by ordinary usage and most tried to

accommodate the diversity of usage by insisting on qualifications. The following definition of "quality of life" provides a good illustration of the whole set of proposed definitions.

> *Quality of life* usually refers to the degree to which a person's life is desirable versus undesirable, often with an emphasis on external components, such as environmental factors and income. In contrast to subjective well-being, which is based on subjective experience, quality of life is often expressed as more 'objective' and describes the circumstances of a person's life rather than his or her reaction to those circumstances. However, some scholars define quality of life more broadly, to include not only the quality of life circumstances, but also the person's perceptions, thoughts, feelings and reactions to those circumstances. Indexes that combine objective and subjective measures, such as happy life years and healthy life expectancy have also been proposed. (Diener, 2005, pp. 401-402)

The logo for the Canadian Index of Wellbeing says "Measuring what matters". The subtitle suggests a broad understanding of wellbeing as reflected in the last two sentences of the previous quotation. We assume that "overall wellbeing" is roughly synonymous with "quality of life". While the quantity of our lives is notoriously limited to one per person, its quality is as varied as the perspectives or domains from which it is viewed. Viewed from one perspective, a person may be well off, but from another not at all well off. A poor person might have good family relations and spiritual fulfillment while living in a rough neighbourhood with substandard housing. Someone with a good job may suffer from a long and lonely commute every day. In the context of research on wellbeing or quality of life, "measuring what matters" implies measuring our lives from the perspectives that are most important to us. The breadth of our concerns and our methodological procedures for ensuring that they will be adequately addressed will become clearer as we proceed with our story.

## Approaches to Indicator and Index Development

Broadly speaking, one may distinguish three relatively ideal types of approaches to the development of indicators and indexes of wellbeing, each beginning from a different strategic point of departure, but never entirely independent of the others. We may name and characterize them as: (1) *Top-Down*, where one begins by constructing a conceptual scheme of some sort describing one's understanding of wellbeing, including its constituents and determinants; (2) *Bottom-Up*, where one begins by exploring the great variety of available data that might be relevant to most people's understanding of wellbeing; and (3) *Bi-Directional*, where one begins by constructing and exploring somewhat simultaneously; that is, one begins by building a framework and at the same time exploring available data sets for items that could populate the framework.

One might characterize the Top-Down approach as theoretical, the Bottom-Up approach as empirical, and the Bi-Directional approach as pragmatic. Of these three approaches, it is fair to say that the development of the CIW has been and will probably remain pragmatic. Practically

speaking, that means that we proceed patiently, transparently, and flexibly, testing any ideas presented both against the hard evidence yielded by empirical research and against the common sense of the Working Group and as broad a constituency beyond it as our resources allow.

The Working Group settled fairly comfortably and early in its deliberations on the idea that most of the phenomena relevant to human wellbeing or the quality of life at the present time could be conceptualized from seven perspectives or domains – living standards, healthy populace, time allocation, ecosystem health, educated populace, community vitality, and good governance. A separate domain on arts and culture was added later, and some of the original names and content of domains were changed as work progressed. For example, the domain of "good governance" was reconsidered as a domain of "democratic engagement" and that of "arts and culture" changed to "leisure and culture".

It did not require much imagination to come up with such a list. In the late eighth and early seventh century BCE, the poet Hesiod of Ascra wrote the following passages in *Works and Days*:

> Those who give straight judgments to foreigners
> and citizens and do not step at all aside from justice
> have a flourishing city and the people prosper in it.
> There is Peace, the nurse of children, throughout the land,
> and wide-seeing Zeus never ordains harsh war for them.
> Famine and Disaster never attend men of straight judgment,
> but with good cheer they feed on the fruits of their labors.
> For these the earth bears the means of life in abundance …
> But for those who have thoughts of evil violence and
> cruel deeds, wide-seeing Zeus son of Kronos has ordained justice.
> Often indeed the entire city of an evil man suffers, …
> Famine and Disease together, and the people perish.
> Women do not give birth, but houses are diminished … (McKirahan, 1994, p.14)

Familiar themes of a good life are cited in these lines – flourishing and prosperous communities, populated by honest people, living in peace, enjoying the fruits of their labour, without worries about where the next meal will come from, with an absence of disease, and with justice for all. Later in the same poem, Hesiod describes the antithesis of a good life through a kind of inversion of these themes.

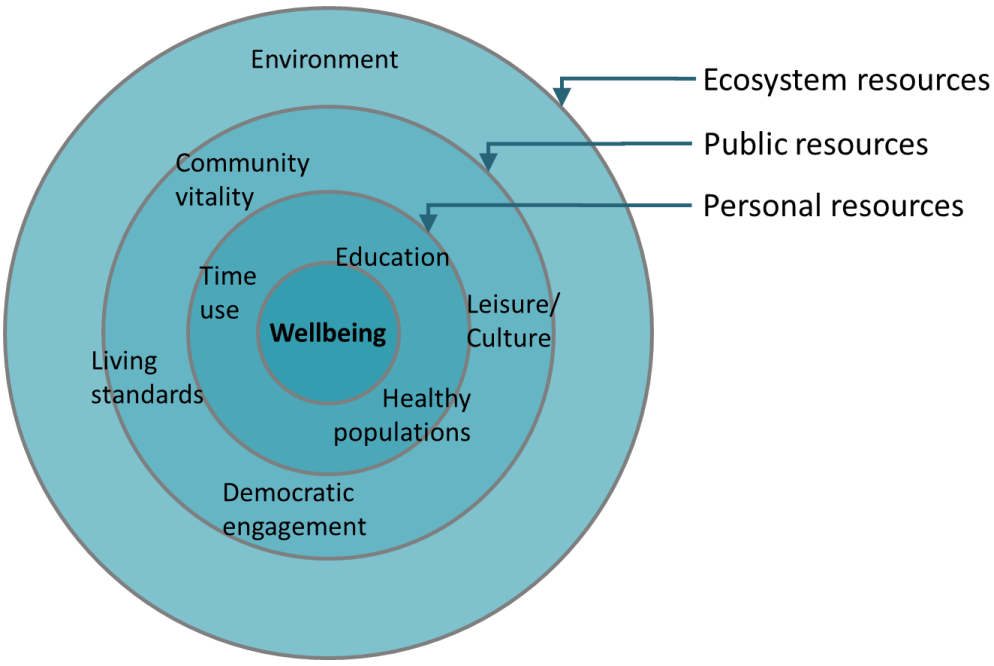
Cummins (1996) scanned 1,500 articles providing data on life satisfaction, looking for "different terms that had been used to describe *domains* of life satisfaction". For an article's terms to be used, the article had to have at least three domains purporting to "represent a broad indication of life quality", and a detailed description of the scales used and average scores obtained for each domain. Oddly enough, "responses to criteria of happiness were excluded", despite the fact that measures of life satisfaction and happiness usually are fairly highly correlated (Michalos, 1991, 2003). Altogether, Cummins found 32 studies meeting his criteria, and those studies used 351 different domain names. The 68 samples described in those studies were "of

four broad types: general population probability or quota samples, general population samples based on a variety of specific criteria, samples of people with chronic medical conditions, and samples comprising people with a chronic psychiatric impairment".

His first aim was to determine how many domain names could be categorized under one of the seven domain headings of his *Comprehensive Quality of Life Scale* (ComQol). The latter's domains included material wellbeing, health, productivity, intimacy, safety, community, and emotional wellbeing. He found that 83% of the 351 domain names could be reasonably classified into one or another of ComQol's seven domains. For example, ComQol's category of "intimacy" includes things like family life, family relations, friendships, marriage, living partner and spouse. He listed the 56 domain names that did not fit into the ComQol categories and concluded that "the question as to the appropriate number of domains must remain open". Nevertheless, he expressed reservations about adding the particular domain of government because "Not only would the inclusion of such a domain exert a strong negative bias on life satisfaction measurement, but also people generally report these aspects of life to be unimportant to them personally". (More on this topic may be found in Michalos, 2005).

Since the game of linguistic construction is relatively open, one can load most general terms with a wide variety of particular ideas. So it is wise to take a fairly pragmatic approach both to the selection of domain names and specific designations. The figure below roughly illustrates the general shape of relationships among the domains of our constructed universe. We call it a *Mandala of Wellbeing* because its circular form is familiar to people in many cultures around the world including, for example, the medicine wheel used by aboriginal peoples and the wheel of life in the Buddhist tradition.

**The Mandala of Wellbeing**



Wellbeing is at the core of the mandala, as the unifying concept of our efforts to measure how Canadian society and individual Canadians are progressing. The core is surrounded by three concentric circles that symbolize the resources we draw upon for our wellbeing. Each of the resources identified in these circles is a domain of the CIW system. The personal resources for wellbeing in the first concentric circle forms the core – healthy populations, education, and time use – are the resources that each of us needs to manage our personal wellbeing. The public resources for wellbeing in the second concentric circle – living standards, community vitality, leisure and culture, and democratic engagement – are the resources we draw upon from the public domain in which we live, from our local communities to the broader society. The ecosystem resources for wellbeing in the outer circle – environment – encompass and affect all of the other circles and domains. We rely on the environment for the natural resources required to sustain human wellbeing in all its manifestations as measured in the domains of the CIW system and summarized in the CIW itself. Taken as a whole, it is safe to say that there is interaction among all the circles, although we are far from knowing all the constituents and determinants of the entire system.

It also is worthwhile to remember that many things are constituents *and* determinants. As Aristotle observed about 2500 years ago, health is good in itself and a means to many necessary and enjoyable ways of being and acting (Michalos, 2011). The Commission on Social Determinants of Health (2007) made the same point saying, "While we see health as having intrinsic value – health as an end in itself – the Commission also recognizes its instrumentality" (p. 10). The satisfaction that one gets from listening to music or playing a musical instrument can be a cause as well as an effect of listening or playing (Michalos & Kahlke, 2010). Dewey (1939) spoke of "the fallacy involved in the position that ends have value independent of appraisal of means involved and independent of their own further causal efficacy" (pp. 42-43). Sen (1999) remarked in many places that freedom has these characteristics. For example, he asserted that "…the expansion of freedom is viewed [by him] as both (1) the *primary end* and (2) the *principal means* of development" (p. 36). As well, he correctly observed that "Such processes as participation in political decisions and social choice cannot be seen as being – at best – among the *means* to development ... but have to be understood as constitutive parts of the *ends* of development in themselves" (Sen, 1999, p. 290). Thus, those who insist upon measuring things that matter *only* as ends in themselves are bound to have an oversimplified view of the world in which we live.

It also is worth noting that the fact that many things that matter are means *and* ends regarded from different perspectives creates special problems connected to the task of crafting measures of sustainable wellbeing or quality of life. In order to know if the current quality of life is sustainable, at a minimum, one must have measures that capture the main constituents of that sort of life (the dependent variable[s], e.g., the CIW) and their rates of change for some determinant period of time, and one must have measures of the determinants (independent, predictor and/or explanatory variables) of that sort of life and their rates of change for the same period. In short, one has to measure whatever it is that one wants to sustain, what resources are required to sustain it or them, and the rates of change of each relative to the

other. While it is true that it can be useful to know the degree to which, for example, health or happiness at one point at time influences health or happiness at a later period (Lucas, 2005; Michalos & Kahlke, 2010b), such information leaves questions about the external or exogenous determinants of health and happiness completely unanswered. Generally speaking, anyone asking a question like, "Why is Jones healthy or happy now?" would hardly be satisfied with the answer "Because he was healthy or happy yesterday". Unfortunately, questions about both endogenous and exogenous determinants of the status quo are vitally important for sustainability analyses. Fortunately, for us anyhow, the focus of attention in this paper is the construction of the dependent variable of such analyses.



## Indicator and Index Selection Criteria and Critical Issues

Having adopted a Bi-Directional approach with eight domains of interest, we proceeded to draft a short list of desirable properties for any acceptable indicator or composite index, i.e., *Acceptability Criteria*. Briefly, we proposed that any acceptable indicator or index of wellbeing should be a statistical measure that has the following characteristics.

### *Acceptability Criteria*

1. Relevant to the concerns of our main target audiences
2. Easy to understand
3. Reliable, valid, and sensitive to changes
4. Politically unbiased
5. Timely, easy to obtain, and periodically updated
6. Comparable across jurisdictions and groups
7. Objective or subjective
8. Positive or negative
9. A constituent or determinant of wellbeing, or both
10. Attributable to individuals or groups of animate or inanimate objects
11. Obtained through an open, transparent, and democratic consultative review process
12. Going to contribute to a coherent and comprehensive view of the wellbeing of Canadians

Acceptability Criteria are not usually specified with great precision (e.g., see Hagerty *et al.*, 2001; IISD, 2009), but they provide useful guidelines for discussions and negotiations over particular indicators and indexes. Good judgement is required in their application in order to prevent the achievement of some goals at the expense of others. For example, measures that

are easy to obtain and update must not be allowed to arrest the development of new measures that might have greater validity and make a greater contribution to a comprehensive view of wellbeing.

Additional complications arise when one considers the array of *Critical Issues* that have to be settled in order to assemble a set of indicators or indexes satisfying the final criterion in the list (i.e., contributing to a coherent and comprehensive set). Inspection of the following list reveals that our Working Group settled some of the Critical Issues in the very specification of our Acceptability Criteria. Explicitly or implicitly, every indicator or index must be specified by addressing the following issues and selecting available options that satisfy the Acceptability Criteria.



## *Critical Issues*

1. *Individual, group, or both bases*: e.g., per capita incomes are inferred attributes applying to individuals, while unemployment rates are inferred attributes applying to groups.
2. *Spatial coordinates*: e.g., the best size to understand air pollution may be different from the best size to understand crime.
3. *Temporal coordinates*: e.g., the optimal duration to understand resource depletion may be different from the optimal duration to understand the impact of sanitation changes.
4. *Population composition*: e.g., analyses by language, sex, age, education, ethnic background, income, and so on, may reveal or conceal different things.
5. *Domains of life composition*: e.g., different domains like health, job, family life, housing, and so on, give different views and suggest different agendas for action.
6. *Objective versus subjective indicators*: e.g., relatively subjective appraisals of housing and neighbourhoods by actual dwellers may be very different from relatively objective appraisals by independent “experts”.
7. *Positive versus negative indicators*: negative indicators seem to be easier to craft for some domains, which may create a biased assessment, e.g., in the health domain, measures of morbidity and mortality may crowd out positive measures of wellbeing.
8. *Input versus output indicators*: e.g., expenditures on teachers and school facilities may give a very different view of the quality of an education system from that based on student performance on standardized tests, and both may be very different from assessing whether the population at large is becoming more literate, knowledgeable, educated, and wise.

9. *Benefits and costs*: different measures of value or worth yield different overall evaluations as well as different evaluations for different people, e.g., the market value of child care is far below the personal, social, or human value of having children well cared for.
10. *Recipient populations*: who should be included as a recipient for particular benefits and costs?
11. *Measurement scales*: e.g., different measures of wellbeing provide different views of people's wellbeing and relate differently to other measures.
12. *Research personnel*: e.g., different stakeholders often have very different views about what is important to monitor and how to evaluate whatever is monitored.
13. *Report readers*: e.g., different target audiences need different reporting media and/or formats.
14. *Aggregation function*: e.g., once indicators are selected, they must be combined, integrated, or aggregated somehow in order to get a coherent story or view.
15. *Distributions*: e.g., because measures of central tendency (i.e., means, medians, and modes) can conceal extraordinary and perhaps unacceptable variation, choices must be made about appropriate representations of distributions.
16. *Distance impacts*: e.g., people living in one place may access facilities (e.g., hospitals, schools, theatres, museums, libraries) in many other places at varying distances from their place of residence.
17. *Causal relations*: prior to intervention, one must know what causes what and what might be jointly interacting with what, which requires relatively mainstream scientific research, which may not be available yet. If possible, correlations among variables should be explored with a view to discovering possible evidence of dependence or independence, redundancy and double-counting.
18. *Discount rates*: how much should one discount costs and benefits delivered sometime in the future compared to those delivered today?
19. *Confidence levels*: what levels of confidence should one require to accept any particular claim or measure?
20. *Validators or Auditors*: who should decide if any assessments are adequate or appropriate?
21. *Validating or Auditing criteria*: what criteria should be used to assess the adequacy of validators' or auditors' assessments, the adequacy of the procedures used for validation, and even the adequacy of the answers to questions raised with the previous 20 issues?

The last question reveals the threat of an infinite regress, a circular argument, or an arbitrary end to analysis. Clearly, none of these options is very attractive, but it is in the very nature of foundational work that such a point must be reached. In any event, supposing that one had only two alternatives for each of the 21 Critical Issues (an absurdly conservative supposition), at least 2,097,152 different sets of indicators and/or indexes might be constructed. If nothing else, this suggests the size of the working space for indicator and index development.

## *Advantages of a Composite Index*

While awareness of the need for good judgement with the help of 12 Acceptability Criteria and sensitivity to 21 Critical Issues provides important basic, relatively generic information for distinguishing useful from useless indicators and indexes, there are additional considerations with specific reference to composite indexes. There is a substantial body of literature devoted to the question of whether, all things considered, an array or profile of diversely related individual indicators of wellbeing in distinct silos is more useful than a single composite index. The Stiglitz Commission (2009, pp. 16-17) claimed that "when driving a car, a meter that added up in one single number the current speed of the vehicle and the remaining level of gasoline would not be any help to the driver. Both pieces of information are critical and need to be displayed in distinct, clearly visible areas of the dashboard". Fair enough. However, most people would rather have a slice of cake than five separate dishes of precisely measured eggs, flour, sugar, butter, and milk. Besides looking for models in machinery with discrete parts, wellbeing researchers should be looking at the arts of cooking, weaving, painting, music, and literature where holistic thinking and orchestrated designs are known to produce qualities (including information) not present in their distinct parts.

We unhesitatingly assert that individual indicators and composite indexes, which are by no means mutually exclusive, each have advantages and disadvantages that vary with the particular uses to which they are put, as well as with the way they are put, in what circumstances, at what time, with what resources and constraints, and by whom, to mention only a few of the most obvious conditions for successful or unsuccessful applications. Setting aside such essentials, in this and the next section we present some of the most salient advantages and disadvantages of using composite indexes, since that is the main focus of this paper. All of the purported advantages and disadvantages are listed without positive or negative comments. The fact that we have constructed and are presenting a particular composite index should be taken as evidence that we find the purported advantages of such indexes attractive enough to warrant serious consideration and exploration of their feasibility for the CIW. However, we are sensitive to the disadvantages as well as the advantages, and we believe our readers should have access to the same material. Taken together, the Acceptability Criteria, Critical Issues, and purported advantages and disadvantages should help readers assess the CIW through a fairly sophisticated lens. In broad strokes, readers should be well prepared to understand many of the complexities involved, have an idea of what to expect and how best to use such an index. (Interested readers may want to match our discussion against the guidelines offered by Ravallion (2010) for developing reasonable "mashup indexes").

At the broadest level of consideration, and changing our single-tree-representing-a-forest metaphor, it is helpful to think of our composite index as merely the door or point of entry into the whole set of indicators and indexes constituting the CIW System. Just as the entrance to a building should be attractive, but not exhaustive of the building's distinctive features and qualities, our composite index should provide an initially appealing introduction to our comprehensive set of measures of things that matter most to Canadians.

The following list of advantages of composite indexes was constructed from Saltelli (2007), Nardo *et al*. (2005), Booysen (2002), and Michalos (1980).

1. A single composite index yielding a single numerical value is an excellent communications tool for use with practically any constituency, including the news media, general public, and elected and unelected key decision-makers.
2. Such indexes provide simple targets facilitating the focus of attention.
3. The simplicity of a composite index facilitates necessary negotiations about its practical value and usefulness.
4. Reduced transaction costs of negotiations with such indicators increase the latter's efficiency and effectiveness, probably leading to the development of better policies and programs.
5. Such indexes provide a means for simplifying complex, multi-dimensional phenomena and measures.
6. They make it easier to measure and visually represent overall trends in several distinct indicators over time and/or across geographic regions and/or population groups.
7. Increases in the ease of measuring and representing trends increases our ability to predict and possibly manage future trends.
8. They provide a means of comparing diverse phenomena and assessing their relative importance, status or standing on the basis of some common scale of measurement, across time and space.
9. Increases in the comparability of phenomena lead to increases in the capacity to make holistic assessments and balanced judgements about them.
10. Increases in the capacity to make such holistic assessments and judgements reduce the likelihood of a public agenda being unduly influenced by relatively narrow interests of a few at the expense of broader interests of many.

11. Because they require construction based on conventions agreed upon by potential users, inventors have considerable flexibility for including desired and excluding undesired features.
12. Because the aim is to construct comprehensive indexes ranging over diverse phenomena, researchers will tend to cast their exploratory resources and conceptual nets broadly, leading to greater collaboration among disciplines and richer explanatory scientific theories.



## *Disadvantages of a Composite Index*

In a fine review article, Booysen (2002) presented 15 problems connected to the construction of composite indexes, although he clearly recognized significant advantages. Briefly, here is his list, supplemented with some suggestions from Nardo *et al*. (2005), Saltelli (2007), and Michalos (1980).

1. A single index must oversimplify complex issues.
2. A single index requires all issues to be significantly comparable.
3. Oversimplified messages will give misleading policy directions, leading to poor policies and programs.
4. Oversimplified measures encourage invidious comparisons among communities, provinces/states, nations, and regions.
5. There will be an ad hoc selection of domains, variables, weighting, and aggregation functions.
6. Ad hoc selections will increase the influence of statisticians and technically trained people at the expense of democratically elected representatives and ordinary citizens.
7. There will still be politically motivated, biased selections.
8. Redundant variables and double-counting will occur.
9. Particular issues will be buried in composite figures, including changes in component variables that significantly increase or decrease the composite figures.
10. Variation and inequalities will be buried in average figures.
11. GDP per capita contains as much information as any alternative composite.
12. If an alternative composite is found, it will lead to the same sort of group-think that surrounds GDP.
13. Index values have no clear meaning.
14. Values of domains, variables, and indexes vary over time.

15. Ends and means will be improperly mixed.
16. Composite figures lack practical value, resulting from all their difficulties.
17. Worse, the search for composite measures may lead to political paralysis while the search goes on.

We will return to some of these issues later when we comment on specific advantages and disadvantages of the CIW.



## Validating the CIW: Basic Concepts

Having negotiated our way to a selection of a set of eight broad domains of interest, 12 Acceptability Criteria, 21 Critical Issues, and a commitment to build a composite index as a simple gateway to deeper exploration, analysis, and understanding, it is necessary to construct some procedures for validating the index. Validity is included in our list of Acceptability Criteria and validity issues permeate all scientific research and a great deal of ordinary life under diverse names (e.g., issues concerning the truth of declarative sentences) and involve similar problems. The issues are so important to our field that a special volume (506 pages) of the journal *Social Indicators Research* was published in 1998, edited by Bruno Zumbo and called "Validity theory and the methods used in validation: Perspectives from social and behavioral sciences". In the Introduction to that volume, Zumbo (1998a, p.1) wrote,

> The concept, method and process of validation are central to quality-of-life and social indicators research, for without validation, any inferences made from a measure are meaningless … lest we fall into traditional camps, validity applies equally to both so-called subjective and objective indicators. The issue at hand is that one needs to make an inference from a score about the state or status of an observational unit whether it is something that is at first glance objective (e.g., annual earnings), or subjective (e.g., self-reported well-being).

Validity issues concern relationships between scores of some sort and real things scored, whether the real things are students enrolled in classes, or someone's perceptions, beliefs, feelings, and evaluations of those students. Generally speaking, a scale, measure, indicator or index is said to be *valid* (in statistical terms) to the degree that it accurately measures what it is supposed to measure. Linguistic usage is not entirely settled, but at least seven types of validity evidence may be distinguished. For our purposes, it is sufficient to describe four commonly encountered sources of validity evidence.

A scale is said to have *face-validity* to the extent that it fairly and transparently appears to directly measure what it is supposed to measure. Although it is debated in the literature whether face-validity is a useful source of validity evidence, because we are often assessing subjective experiences, some attention must be given to this sort of validity evidence. For example, the most frequently used standard measures of self-reported happiness have good

face-validity. The 3-step (i.e., 3 response categories) happiness scale of Gurin, Veroff and Feld (1960) simply asks, "Taking all things together, how would you say things are these days – would you say you are very happy, pretty happy, or not too happy these days?" Scales using 5, 7, 10, or even 11 response categories are often used, with equal face-validity. On the other hand, the 23-item *Memorial University of Newfoundland Scale of Happiness* (Kozma & Stones, 1980) is more complicated and has rather less face-validity.

A scale is said to have *criterion-related validity* insofar as it has a significant correlation with some other relevant measure or criterion. To take a simple example, the written driver's test one must pass in order to take a road test is criterion-related valid exactly insofar as success on the former is correlated with success on the latter. In principle, all of the items in our list of acceptability criteria could be used for assessing criterion-related validity, depending on what measures happen to be available, what levels of technical sophistication researchers or interested target audiences have, what purposes people have for their measures, and other context-specific matters. Since many of the indicators and indexes included in the CIW have been used extensively for many years (e.g., self-reported health, smoking behaviour), they have been validated in various ways. The frequently used 5-step scale of self-reported health (from "poor" to "excellent") has good face-validity as well as substantial criterion-related validity with a variety of measures of mortality and morbidity (CDC, 2000; Idler & Benyamini 1997).

A scale is said to have *content-validity* insofar as it adequately or completely refers to the relevant content of some area or domain to be measured. There are no generally accepted criteria for establishing this sort of validity (Nunnally, 1967; Zeller & Carmines, 1980). Philosophers of science distinguish ordinary or pre-scientific concepts to be defined from the scientific definitions themselves, insisting that one should not attempt to construct scientific definitions until there is adequate clarification of the pre-scientific ideas (Michalos, 1971). In our field, as explained earlier, different people have different ideas or preconceptions about happiness, health, wellbeing, and quality of life, and it is important to clarify these ideas before attempting to construct a generally acceptable scientific definition of any of them. Researchers constructing the CIW are engaged in the very complicated task of building a scientifically acceptable, measurable index of wellbeing that sufficiently includes enough of the content of most Canadians' ideas and preconceptions about overall wellbeing, the quality of our lives, and what really matters to us, that the index will become generally accepted.

A scale is said to have *construct-validity*, in one sense of the term, insofar as it is appropriately correlated to other scales that are theoretically related to it. For example, theoretically, one would suppose that people who are typically depressed would not be experiencing a high degree of life satisfaction. Hence, one would expect that construct-valid life satisfaction scores would be negatively correlated with the scores from any valid scale measuring depression (e.g., with the *Beck Depression Inventory*). Michalos (1991) reviewed a few dozen highly correlated theoretical constructs showing the construct-validity of several standard measures of happiness and life satisfaction.

In another sense of the term, a scale is said to have *construct-validity* insofar as the observed scores obtained from it reflect the underlying theoretical construct the scale is intended to measure. The determination of this sort of validity involves the use of sophisticated structural equation models, but some early results relevant to our project are very promising. Andrews (1984) produced a detailed study of a wide variety of features of survey research items used in subjective wellbeing studies. Among other things, he measured the effects of 13 aspects of survey design, including the number of answer categories in a response scale, the presence of a "Don't know" option, category labeling, explicit midpoint, absolute versus comparative items, length of item introductions and items themselves, numbers of items in a scale, position of particular items, data collection procedures, social desirability, content specificity, respondent experience, and content salience. His general conclusion was that "a typical survey item, when administered by a respected survey organization to a general population sample, can be expected to yield 50-83% valid variance, 0-7% method effects variance, and 14-48% residual variance … over two-thirds of the variation in measurement quality could be explained by 13 survey design characteristics" (Andrews, 1984, p. 409). In brief, his results indicate that at our best, those who are trying to measure individual subject-perceived wellbeing with a variety of the most frequently used standardized satisfaction or happiness scales are probably measuring a substantial chunk of precisely that.

The validation of the CIW is situated within a traditional framework of validity (e.g., Cronbach, 1971; Cronbach & Meehl, 1955) that is widely used in the social indicators and quality of life literature. As evidenced as early as Zumbo's 1998 special issue, there is a move in the literature and in validation practices to embrace aspects of recent developments in validity theory (e.g., Kane, 2006; Messick, 1989; Shepard, 1993; Zumbo, 2007, 2009). Validity has become an increasingly expansive concept, moving from distinct "types" of validity that could be demonstrated through a single correlation coefficient or factor analysis to more nuanced theories that advocate using sociological and contextual factors as evidence (e.g., Zumbo, 2007, 2009). In this view, validity is no longer seen as a static property of tests, scales, or indexes, but rather as an integrated judgement about the degree of justifiability of inferences we make based on such measures (Messick, 1989). As well, a multilevel view of validity has been recently developed that is meant to inform social indicators practice and particularly to address typical uses of social indicators in complex multilevel social systems (Zumbo & Forer, 2011). There is a recognized gap between validity theory and practice, although the exact nature and mechanism of this gap remains unclear (Messick, 1988; Wolming & Wikstrom, 2010). As the idea of validity has become more expansive, it also has become more complex, giving rise to debates in the field about what sorts of evidence are needed in different contexts and how best to synthesize the evidence.

## *Controversial Issues Concerning Validity*

Some researchers today (e.g., Cummins, 1996) assume that one or another of the frequently used measures of happiness or life satisfaction can be used as criterion variables (i.e., gold standards) to assess overall human wellbeing. Such indicators essentially measure individuals' experiences, perceptions, or reports of experiences, and the psychometric properties of these

indicators are well-known. There are over 40 years of research (Michalos, 2005) showing that these measures are largely unaffected by the great issues of our time like the state of the natural environment (i.e., air, water, and land pollution), the sustainability of production and consumption patterns, the earth's carrying capacity for waste products, other people's wars, poverty, disease, criminal victimization, ignorance or just plain bad luck, as well as the arts and culture that reveal some of humanity's greatest achievements. On the other hand, these measures are affected by such things as survey design features (explained below), individuals' temperament (genetic hard-wiring), transient moods, culture, community features (communities of place and interest), seasons, weather, social support, income, life events (e.g., births and deaths, job loss), individuals' perceived discrepancies between the real conditions of their lives and desired conditions, upward and downward comparisons with discrepancies between the conditions of their own lives versus those of selected reference groups, perceived discrepancies between what they have and think they deserve (equity issues), and comparisons between their past experiences and/or imagined future experiences (Diener, Sapyta, & Suh, 1998; Diener & Seligman, 2004; Michalos, 2005, 2003).

While everyone working on the CIW accepts the importance of including measures of personal experiences in any reasonable index of overall individual or community wellbeing, using correlations of measures of happiness or life satisfaction with the CIW as means of validating the latter would not be feasible for several reasons. Measures of happiness or life satisfaction should be included as constituents of human wellbeing (although they are absent in the current version of the CIW), and as constituent elements of the CIW, any correlations between the former and the latter would involve some autocorrelation (i.e., correlating something with itself). The same problem would arise if, for example, one used poverty rates, health adjusted life expectancy, or any other constituent element of the CIW as a criterion variable. Besides this technicality, such a procedure would be tantamount to evaluating a measure of overall wellbeing by how well it correlates with a measure of a relatively small part of such wellbeing (individual perceptions of their own happiness or life satisfaction). Experienced or perceived wellbeing, sometimes called subjective wellbeing, cannot be a validating criterion for the CIW because much of wellbeing is not salient or experienced. Experienced wellbeing is necessary, but not sufficient for overall human wellbeing. We must know how experienced wellbeing is related to overall wellbeing, but the former (one part) must not be confused with the latter (the whole). The CIW is intended to measure wellbeing as a whole (see also Diener, 1994; Kahn & Juster, 2002).

Two reasons have been most frequently offered for using measures of happiness or life satisfaction obtained from survey research as validating criteria for a general index of wellbeing, each of which has some problems. First, some people seem to suppose there is something particularly democratic about using such measures. There is a tendency for some researchers to regard survey research results as an indication of "the collective will" of those surveyed. However, survey research is not designed to and cannot serve that purpose. Vague as the idea is in democratic theory, the "collective will" of a community of people is the sort of thing determined or constructed by means of public discussion and debate, according to some more or less explicit rules, including rules of participation and closure (Michalos, 1978). Because

survey researchers usually do not attempt to present a balanced set of relevant considerations to a respondent, to elicit questions for debate, review alternatives and assess probabilities or preferences prior to asking any questions, it is likely that a typical respondent is not giving answers in the light of such considerations. So, those who regard survey research results as indicators of a "collective will" in the sense understood by democratic theorists are misleading themselves and perhaps others. Sten Johansson, a Swedish pioneer in social reporting and quality of life research, has warned us of being misled in this way for many years, most recently in Johansson (2002).

According to Converse (1987):

> Virtually all of the major figures before 1910 – including George Gallup, Elmo Roper, and Archibald Crossley in the private sector, or Henry Wallace and Rensis Likert in the government at the [U.S.] Department of Agriculture – were strong on democratic principles and pleased to provide a means that the voice of the people might be more clearly heard to compete with the few voices in the ears of power. (p. S15)

One would hope that those pioneers recognized the danger in their becoming and reproducing a new elite group of technicians who might stifle and distort the very voices they were trying to amplify. However, regardless of the democratic sensitivities of the pioneers, it is vitally important that all of the researchers most intimately involved in the validation of the CIW recognize the fundamental difference between results of a public opinion poll and a democratically determined collective decision of a community. While we should be open to listening to any voices addressing the CIW and its validation, the voices coming from the free exchange of views among well-informed people in democratic fora should be given greater weight than the aggregated views of randomly selected individuals providing relatively quick answers to interviewers' questions. Sen (1999) was certainly right when he wrote that "one of the strongest arguments in favour of political freedom lies precisely in the opportunity it gives citizens to discuss and debate – and to participate in the selection of – values in the choice of priorities" (p. 30) [i.e., what matters most to them].

A second reason offered for using happiness or life satisfaction scores as criteria in assessing criterion-related validity is that it is assumed that somehow human beings are able to implicitly assemble all considerations that are relevant to their wellbeing, assign appropriate weights to each one, and wisely aggregate all the information to produce a single score (e.g., Hagerty & Land, 2007). As explained above, we have lots of good research showing that individuals do make such judgements about their own happiness or life satisfaction, but such judgements usually do not involve many features of the real world in which people live and do involve many features that are individually and socially constructed more or less wittingly or unwittingly. Some years ago Dawes (1979) made the point quite boldly, saying "People are bad at integrating information from diverse and incomparable sources" (p. 574). While people's assessments of their own wellbeing should be given some privileged status in our overall assessments, such assessments cannot be regarded as merely given and incorrigible. Regarding the virtues of individuals' *implicit* aggregation of diverse bits of information into a single

composite judgement versus a community's *explicit* aggregation of such information into a single composite index, Michalos (2008) wrote that "moving from the Pandora's Box of aggregation problems in the visible world to the invisible Black Box inside people's heads does not strike me as a progressive research program" (p. 360). (Additional problems with using happiness or life satisfaction as criteria can be found in Sen [1985, 1999], while on the other hand, Hagerty and Land [2011] recommend some such criteria.) Any tests of the criterion validity will probably require the construction of instruments designed specifically to elicit judgements about the CIW, its content and methodology.

Because the CIW is intended to capture our ideas of overall wellbeing and not merely subjective wellbeing, we are a long way from having a theoretical model to fit against reality. That is, we are a long way from being able to assess the construct-validity of the CIW in this robust sense. As currently conceived, the CIW will be built upon 64 indicators and indexes in eight domains, covering diverse levels of aggregation. For example, individual households are nested in neighbourhoods, which are nested in towns or cities, which are nested in metropolitan areas within diverse water and/or air sheds, and these are nested in different regions across the vast geography known as Canada (Michalos, 2008). The multi-level modelling required to understand the impact of these different areas is seldom addressed and not well-understood. As well, for any two indicators in the whole system, say **A** and **B**, we often do not know if **A** influences **B**, **B** influences **A**, or the two are mutually influential. For any three indicators in the whole system, say **A**, **B**, and **C**, we hardly know if **A** influences **B** and **B** influences **C**, or if **A** and **B** together influence **C**, or if **A**, **B**, and **C** are together influenced by something prior to **A**. To complicate matters even more, because the influence of any indicator on any other is time dependent, even when discrete time analyses show that, say, **A** influences **B** in a certain period of time, **B** might influence **A** at a later time. For example, parent-child dependency relationships are such that in our early years we are largely dependent on our parents and in our later years our parents may be largely dependent on us. Without continuous time series analyses there is no way to be sure that causal arrows always run in the direction shown by discrete time analyses and no way to know at what point in time the causal arrows may be reversed. Unfortunately, very few social scientists work with continuous time analyses (Oud & Delsing, 2010). In other words, the state of the development of the CIW is still largely at the stage of identifying elements that matter most, the **A**s, **B**s, and **C**s. So we can know relatively little about the functional form of the relationships and/or equations connecting all the elements.

Besides validity, as the third item in our list of acceptability criteria indicates, the CIW must be reliable. A scale is said to be *reliable* (in statistical terms), in one sense, to the degree that it yields similar results from measuring the same phenomena in the same way at different points in time. With multiple-item scales, one may also measure their internal reliability, which is the degree to which their constituent items are consistent (i.e., correlated). For single item scales, reliability is frequently measured by having people respond to the item at different points in time and correlating the responses using zero-order correlations. This is *test-retest reliability*.

Because the constituent indicators and indexes of the CIW are currently drawn from a wide variety of surveys, from a variety and number of points in time (occasionally as few as two),

with a variety of definitions and conceptual frameworks, it is impossible to assess the internal consistency of all the elements of the system (i.e., impossible to do useful, correlational, and multi-correlational analyses for many constituents of the system), although some subsets of the whole system do permit such assessments. However, because some constituent indicators and indexes of the CIW are measures of effects, others are measures of causes, and some are measures of both causes and effects, it should not be assumed that all constituent indicators and indexes should be correlated positively or negatively. They may be independent (Bollen, 1984; Fayers & Hand, 2002; MacCallum & Browne, 1993; Wright & Feinstein, 1992). For example, drinking water quality and fear of walking alone at night might both be important indicators of the quality of a person's life or wellbeing although measured changes in indicators of either variable would not be expected to have an impact on the other. Fayers and Hand (2002, pp. 234-237) explained the differences between cause- and effect-indicators, and the uses of each, very well, although their primary focus was on composite indicators used in the field of health-related quality of life research.

> … psychometricians try to measure *a single attribute with multiple items* and use [a "traditional framework of validity" along the lines described above] … to demonstrate that the multiple component items are all measuring (more or less) the same single attribute (latent variable). Clinicians try to measure *multiple attributes with a single index* and aim their strategies at choosing and suitably emphasizing the most important attributes to be included in the index. (Fayers & Hand, 2002, p. 235)

Just as it should not be assumed that the independence of some variables from others is evidence of irrelevance, it should not be assumed that a high correlation among some variables is evidence of pernicious redundancy. Hagerty and Land (2007) showed that the three variables of the UNDP's Human Development Index (GDP per capita, life expectancy, and education) are highly correlated (across countries) from 0.82 to 0.77, but that index has been enormously successful and useful at drawing the world's attention to the diversity of development and its consequences for people's lives. Besides, nobody would suppose that these high correlations would provide a good reason to, for example, replace GDP per capita with life expectancy in the National Income and Product Accounts or use GDP per capita as a primary dependent variable in national population health reports.

## Validating the CIW: Basic Processes

Kanowitch, Michalos, Slotek, and McKessock (2007) provided a sketch of the historical background of the development of the CIW System. Here it is only necessary to describe the basic processes designed to be used by the Project Management Team and the National Working Group in the development of the reports for each domain and for this report on our general approach to a composite index. These processes have been developed over time, expressed more or less explicitly, and used with more or less success. There are eight fundamental steps.

1. *Expert Literature Reviews and/or Environmental Scans*: For each domain, the Project Management Team commissions a literature review and/or an environmental scan by one or more experts in the field. These reviews are supposed to provide state of the art overviews of relevant research, including especially Canadian research and time-series data availability. There might be as many as 50 or more potentially useful time-series available in any particular domain. Besides providing a comprehensive survey of the "things that matter" in a domain, the commissioned experts are supposed to provide a list of headline indicators or indexes in the domain. Headliners are presented in fairly standardized formats giving readers information on such essentials as data sources, frequency of reporting, availability, relevance, cost, specific wording of items, level of geography, use in other systems, and general considerations regarding advantages and disadvantages. There might be 10 to 20 potential headliners among all available indicators. Finally, to ensure that reviews and results are not limited to available material, commissioned experts are asked to recommend new indicators, domains, procedures, and topics worth considering in the future. These recommendations are presented in Appendix 1 of this paper.

2. *Management Team and Peer Review*: Each commissioned review is circulated among the Project Management Team and five or more experts who write reviews usually recommending a variety of revisions. Different experts are typically required for different domains, as is the case for most peer reviews. Given the centrality of the composite index for the CIW System, a greater number of critiques is warranted and was in fact obtained.

3. *Literature Reviews/Environmental Scans Revised*: Revised reviews are circulated to the Project Management Team and a decision is made to proceed to the next step or get additional revisions.

4. *National Working Group Assessments*: When the Project Management Team decides that the reviews are suitable for discussion by the National Working Group, the reviews are circulated among the group with a voting matrix. The voting matrix includes a list of each proposed headliner with a five-point rating checklist running from 5 = "very important" to 1 = "not at all important" to the composite CIW. Members of the National Working Group are asked to make preliminary, independent assessments of each proposed headliner prior to hearing the full discussion of the group about the proposals and then to make final assessments following the discussion. These assessments are made primarily against the 12 Acceptability Criteria, which technically speaking makes them *content-validity assessments*. Due to scheduling difficulties, this step has not been taken for all domains.

5. *Final Headliner Selection*: The National Working Group's assessments are then forwarded to the Project Management Team for summary

calculations and a final selection of eight headliners is made for inclusion in the composite index for the domain and for the CIW itself.

6. *Domain Index and CIW Construction*: The finally selected set of headliners for each domain is then forwarded to the team of Michalos and Sharpe to construct the Domain-Specific Index and to add it to the composite CIW. The precise steps used to calculate these indexes are described in the following sections of this paper.

7. *Focus Group Reviews*: As described in the *Foundational Document*, focus groups have been organized across the country periodically to obtain additional content-validation for the processes and products of the domain investigations and composite CIW. These meetings with a wide variety of Canadians from all walks of life are an essential feature of our overall validating process. They provide necessary grounding of our project. If most ordinary Canadians cannot hear their voices and see themselves, their interests, and their values in the CIW, then our work will have been in vain and we will have failed. Munda (2004) captured the importance of such reviews when he wrote,

> For the formation of contemporary public policies, it is hard to imagine any viable alternative to extended peer communities…They are called 'citizens' juries', 'focus groups', or 'consensus conferences', or any one of a great variety of names; and their forms and powers are correspondingly varied. But they all have one important element in common: they assess the quality of policy proposals, including the scientific and technical component. And their verdicts all have some degree of moral force and hence political influence. Here the quality is not merely in the verification [validation], but also in the *creation*; as local people can imagine solutions and reformulate problems in ways that the accredited experts, with the best will in the world, do not find natural. (p. 667)

8. *Continuous Validation*: As explained in the previous section, there are diverse meanings to the term "validation" and some kinds of validation are feasible for some indicators and indexes in some domains, but not in others. The same thing may be said for Domain-Specific indexes and the composite CIW itself. Most importantly, it must be recognized that the work of validating indicators and indexes must be continuous. In a paper focused on ethical issues related to statistical modeling, Kleijnen (2001) suggested two useful metaphors concerning validation:

> A car is periodically returned to the garage for maintenance; similarly a model may be returned to its builders, for updating. With other software it is well-known that maintenance is a crucial – and expensive – part of the life cycle! ... Another metaphor is the instructions that come with most medicines: these instructions warn against all kinds of undesirable side-

> effects. Likewise, the documentation of a model should warn against improper usage. And likewise, this documentation should be updated continually. (pp. 226-227)

As the world changes, new issues become salient, new knowledge and technology becomes available, and some of the things that matter most to most people today may not matter most to most people in the future. We saw earlier that a core set of features of a good life identified by ancient Greek philosophers and poets remains as attractive today as it was in 800 BCE, although some things have been properly discarded. The task of validating the CIW and its constituent parts is perhaps best regarded as the latest contribution of some human beings to an initiative that began long before us and will not end until the last human breaths her or his last breath.

## Our Approach to a Composite Index of Wellbeing

Tables 1a, 2a, 3a, 4a, 5a, 6a, 7a, and 8a list eight headliners for the domains of living standards, healthy populations, community vitality, democratic engagement, leisure and culture, time use, education, and environment, respectively. It would be unwieldy to include background information about all 64 headliners in the CIW here. However, detailed descriptions of each indicator and some considerations leading to its selection as a headliner may be found in the background documents covering each domain. Data limitations reduced our ability to use all indicators identified as headliners in all domains.The living standards domain is much richer than the other domains in long term data. Because most of our health statistics were drawn from the National Population Health Surveys which began in 1994, we made 1994 our base year. Since measures of change are significantly influenced by the particular baseline against which changes are measured, the selection of any particular year is not a trivial matter. For example, one may make changes in economic indicators like GDP per capita and unemployment rates look more or less favourable by selecting baselines at appropriate years in business cycles. Our selection of a baseline year was practically entirely determined by our interest in maximizing the number of headliner indicators available for the health domain. Our selection of 2008 as our final year of review was determined by data availability. There were relatively few statistics available for 2009 and fewer still for 2010 for most of our headliners. Having selected 1994 as our baseline year, certain indicators in all domains were immediately lost as a result of insufficient data points; for example, wealth distribution was available only for 1999, food security for 2001, and satisfaction with health care services for 2003.

### *Living Standards*

The headliner indicators used for the *Living Standards Domain* (Table 1a) include four negative indicators (i.e., increases in numerical values indicate decreases in some aspect of wellbeing) and four positive indicators (i.e., increases in numerical values indicate increases in some aspect of wellbeing). The negative indicators are flagged with "**n**" and the positive with "**p**". They are listed here in the order in which they appear in the headings of the columns of Table 1a reading from left to right: (1n) ratio of the top to the bottom quintile, after tax income of economic families, (2p) after tax median income of economic families in 2008 dollars, (3n) incidence of economic families in poverty (i.e., below Statistics Canada's Low Income Cut Off points), (4p) scaled value of economic security index, (5n) incidence of long-term unemployment, (6p) employment rate, (7p) CIBC index of employment quality, and (8n) RBC housing affordability index.

**Table 1a: Trends in Living Standards Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Living Standards[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1n | 2p[b] | 3n[c] | 4p | 5n | 6p | 7p | 8n |
| **1994** | 4.83 | 51,700 | 14.0 | 0.571 | 17.4 | 58.4 | 100.6 | 41.5 |
| **1995** | 4.92 | 51,200 | 14.5 | 0.576 | 16.3 | 58.7 | 101.6 | 39.6 |
| **1996** | 5.18 | 51,100 | 15.2 | 0.564 | 16.3 | 58.5 | 100.1 | 36.7 |
| **1997** | 5.31 | 51,400 | 15.0 | 0.563 | 15.6 | 59.0 | 100.2 | 35.2 |
| **1998** | 5.52 | 53,300 | 13.7 | 0.566 | 13.3 | 59.7 | 100.4 | 34.7 |
| **1999** | 5.44 | 55,100 | 13.0 | 0.570 | 11.3 | 60.6 | 104.3 | 35.5 |
| **2000** | 5.69 | 56,000 | 12.5 | 0.579 | 10.8 | 61.3 | 105.3 | 36.9 |
| **2001** | 5.56 | 58,300 | 11.2 | 0.551 | 9.0 | 61.1 | 105.6 | 34.7 |
| **2002** | 5.63 | 58,200 | 11.6 | 0.512 | 9.2 | 61.7 | 102.8 | 35.2 |
| **2003** | 5.53 | 58,100 | 11.6 | 0.508 | 9.6 | 62.4 | 100.2 | 35.6 |
| **2004** | 5.72 | 58,900 | 11.4 | 0.516 | 9.1 | 62.6 | 99.0 | 36.8 |
| **2005** | 5.62 | 59,900 | 10.8 | 0.518 | 9.2 | 62.5 | 99.1 | 37.4 |
| **2006** | 5.54 | 61,100 | 10.5 | 0.522 | 8.3 | 62.8 | 98.3 | 41.1 |
| **2007** | 5.50 | 63,400 | 9.2 | 0.530 | 7.1 | 63.4 | 97.6 | 44.3 |
| **2008** | 5.61 | 64,100 | 9.4 | 0.521 | 6.7 | 63.5 | 99.8 | 45.0 |

[a] *Key:* 1n = Ratio of top to bottom quintile of economic families (after tax)
2p = After tax median income of economic families (2009$)
3n = Incidence of persons in low income (%)
4p = Scaled value of CSLS economic security
5n = Incidence of long-term unemployment (%)
6p = Employment rate (%)
7p = CIBC index of employment quality (1994 Q1=100)
8n = RBC housing affordability index

[b] Data for after tax median income of economic family is in constant dollars for 2009.

[c] Data for incidence of poverty based on the percentage of persons below the low after-tax income cut-off.

For positive indicators, each of the eight raw indicator scores of Table 1a is converted into an index of percentage change in Table 1b by dividing every raw score in each column by the first score in the column and multiplying the result by 100. For example, the first score for (2p) after tax median income of economic families gives 51,700/51,700 = 1 x 100 =100.0; 51,200/51,700 = .9903 x 100 = 99.0, and so on.

It should be noticed that the replacement of raw data scores by change scores in the interest of obtaining comparability across the set of indicators was made at the expense of a loss of important information for each indicator. For example, the final change score for economic families' after tax median incomes across the 15 years (124.0) indicates that some progress was made, but it fails to indicate anything concerning the size or adequacy of those incomes.

Clearly, the information contained in Table 1a is at least as important as the information in Table 1b.

**Table 1b: Index of Living Standards Indicators for Canada, 1994 to 2008**

| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1n | 2p | 3n | 4p | 5n | 6p | 7p | 8n | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 98.3 | 99.0 | 96.6 | 101.0 | 106.9 | 100.5 | 101.0 | 104.8 | **101.0** |
| **1996** | 93.3 | 98.8 | 92.1 | 98.8 | 106.7 | 100.1 | 99.5 | 113.2 | **100.3** |
| **1997** | 91.0 | 99.4 | 93.3 | 98.6 | 111.5 | 101.0 | 99.5 | 118.0 | **101.6** |
| **1998** | 87.6 | 103.1 | 102.2 | 99.2 | 130.7 | 102.3 | 99.8 | 119.7 | **105.6** |
| **1999** | 88.9 | 106.6 | 107.7 | 99.9 | 155.0 | 103.7 | 103.7 | 117.2 | **110.3** |
| **2000** | 85.0 | 108.3 | 112.0 | 101.4 | 162.1 | 104.9 | 104.6 | 112.5 | **111.4** |
| **2001** | 86.9 | 112.8 | 125.0 | 96.6 | 192.8 | 104.7 | 105.0 | 119.7 | **117.9** |
| **2002** | 85.9 | 112.6 | 120.7 | 89.7 | 189.4 | 105.7 | 102.1 | 118.0 | **115.5** |
| **2003** | 87.4 | 112.4 | 120.7 | 89.0 | 182.3 | 106.9 | 99.6 | 116.8 | **114.4** |
| **2004** | 84.5 | 113.9 | 122.8 | 90.4 | 192.2 | 107.2 | 98.4 | 112.9 | **115.3** |
| **2005** | 86.0 | 115.9 | 129.6 | 90.7 | 190.4 | 107.1 | 98.5 | 111.2 | **116.2** |
| **2006** | 87.3 | 118.2 | 133.3 | 91.5 | 211.0 | 107.5 | 97.7 | 101.2 | **118.5** |
| **2007** | 87.8 | 122.6 | 152.2 | 92.8 | 246.6 | 108.5 | 97.0 | 93.7 | **125.2** |
| **2008** | 86.1 | 124.0 | 148.9 | 91.3 | 260.4 | 108.8 | 99.2 | 92.3 | **126.4** |

[a] *Key:* 1n = Ratio of top to bottom quintile of economic families (after tax)
2p = After tax median income of economic families (2009$)
3n = Incidence of persons in low income (%)
4p = Scaled value of CSLS economic security
5n = Incidence of long-term unemployment (%)
6p = Employment rate (%)
7p = CIBC index of employment quality (1994 Q1=100)
8n = RBC housing affordability index

[b] Average of living standards indicators

The last sentence of the previous paragraph cannot be overemphasized. Besides being compensatory, our percentage change scale provides a measure of the relative improvement or deterioration from a baseline, but does not provide an absolute measure. Numerical values of the scale can tell us how far our indicators have moved on average from the baseline, but they cannot tell us if the status at the baseline or later was good or bad in itself. In some ways, our scale leaves us in a position like that of the ancient Greeks watching a footrace. At the end of the race they could tell if a runner was one or more steps ahead or behind other runners, but lacking a clock, they could not tell how fast anyone was running. Similarly, a doctor at the time could tell if a patient was feverish compared to most other people, but had no way to measure temperature.

This limitation of percentage change scales is directly connected to another, namely, the problem of ceilings and floors. If some raw scores have natural ceilings or floors beyond which it is practically or physically impossible to make any changes, then a percentage change scale could not provide any useful information. For example, if water quality in some water shed cannot be purified to the point of being drinkable given current technology and other resources, percentage changes from the current status of a water quality indicator will not increase (indicating some improvement). In such cases, the indicator should be withdrawn from the CIW. At this point in time, we do not know if any of our indicators are at or near natural ceilings or floors.

In order to standardize the index values so that increases and decreases in figures uniformly represent improvement or deterioration in wellbeing, respectively, we transformed the values of negative indicators into their reciprocals and turned the latter into percentages. For example, the index value for the first year of the negative indicator (1n) is transformed from 4.83/1 to 1/ 4.83 = .2070, dividing that number by itself gives .2070/.2070 = 1 x 100 = 100.0; the same operation performed for the second value, 4.92, gives us 1/4.92 = .2033, and dividing by .2070 gives us .2033/.2070 = .9821 x 100 = 98.2, and so on. This transformation is non-linear and may distort some trends, but it seemed to be our best option. Reviews of scaling options and problems may be found in Maggino and Zumbo (2011), Nardo *et al*. (2005), and Horn (1993).

Effectively, the technical problem of constructing a unidimensional scale to reasonably represent a multidimensional construct of human wellbeing is solved by creating a *mean of percentage change rate ratios scale* (*percentage change scale* for short*)* with the property just described. Because percentage change scales allow trade-offs between deteriorations on some indicators to be compensated by improvements in others, they may be regarded as *compensatory* scales (Munda, 2006; Munda & Nardo, 2005). While it is true, as Munda has argued in several places, that certain losses (e.g., losing a loved one or all one's drinkable water) might never be compensated by any gains on any other indicator, we spend much of our lives making trade-offs among many things. For example, we might trade off losses of time spent in pleasurable leisure activities with family and friends against gains in income, education, or the pursuit of public goods. The frequency and importance of such trade-offs for living a good life seems sufficient to require the CIW to be compensatory.

Fine definitions and analyses of problems related to compensatory versus non-compensatory scales, strong and weak comparability among scales, and the importance of value-pluralism for multi-criteria evaluation may be found in Martinez-Alier, Munda, and O'Neill (1998). Similar problems have been addressed by ecological economists in discussions of weak and strong sustainability. For example, according to Victor, Hanna and Kubursi (1995):

> The crux of weak sustainability is the assumption that manufactured and natural capital are close substitutes. [while] … The concept of strong sustainability is based on a denial of the degree of substitution that weak sustainability assumes, at least for some classes

> of natural capital…There is no substitute for drinkable water, breathable air or the genetic information that produces the great diversity of species and individuals within the global ecosystem. These components are essential to sustaining life. Contaminated water and air may be made pure but air is not a substitute for water nor for genetic material. All are essential requisites … the real threats to sustainable development stem from the likelihood that some functions of natural capital are indispensable and non-substitutable. (pp. 78, 89)

Although the current form of the CIW is compensatory and is consistent with the assumption of weak sustainability, it is certainly not cast in stone.

The final column of Table 1b lists the figures for a CIW for the living standards domain for each of the 15 years in the 1994 to 2008 period, and the CIW for the domain for the whole 15 year period is the last figure in that column (i.e., 126.4). In short, over the 15 year period, there was an improvement of 26.4% in the living standards domain.

In considering the assignment of weights in a principled way to each headline indicator, we followed a variant of the eighteenth century mathematician Pierre Simon de Laplace's *Principle of Nonsufficient Reason* (Michalos, 1969) for assigning initial probability values to a set of events. The principle suggests that in the absence of a sufficient reason to regard any particular indicator as more important than any other, each indicator should be assigned an equal weight. As suggested above in our comments about compensatory scales, there are many reasons for regarding one or another indicator as more important in some way or other, but what is missing is a good reason for assigning any particular indicator a particular numerical value greater or less than that of some or all other indicators. It is the absence of such a reason that justifies the equal treatment of all indicators here. With the greater understanding of the relationships among all indicators that is bound to come as development of the CIW proceeds, sufficient reasons for diverse weights may appear. Of course, anyone is free to adopt any weights that may appear to be particularly compelling, but the greater the variety of weights assigned to components of the CIW, the less useful the index will be as a common or generally accepted measure of wellbeing. (On the importance of simplicity in the assignment of weights and all other features of composite indexes, see Cox, Fitzpatrick, Fletcher, Gore, Spiegelhalter, & Jones, 1992).

Although we assigned equal importance weights to each headliner following the line of reasoning of classical probability theorists beginning with Laplace, Hagerty, and Land (2007) undertook a more rigorous analysis to try to show "how to construct summary indices (e.g., quality-of-life [QOL] indices) for a social unit [e.g., a community, city, country] that will be endorsed by a majority of its citizens." The conclusions that they reached add support to our position, provided that one accepts their basic assumptions, particularly about the relative importance of survey research for making public policy decisions. For present purposes, it is enough to record their main conclusions and invite interested readers to go back to the source to see their detailed arguments. Most importantly, Hagerty and Land (2007) showed that it is:

… possible to create a QOL index that a majority of individuals endorse (i.e., they agree at least with the direction of the QOL index). Specifically,

1. When correlations among social indicators are all positive …, then agreement will be high regardless of the variation in weights. This highlights the paradoxical result that people may argue *in theory* about whose weights are more ethically appropriate, but *in practice*, their conflicting weights will yield substantial agreement on the overall QOL index …

2. When some correlations among social indicators are negative…, our results are the first to show that disagreement is much rarer than expected and occurs only when the distribution of individuals' weights is (a) bimodal and (b) negatively correlated, i.e., when individuals' weights are diametrically opposed … highly polarized and emotional issues, such as abortion they are more likely to show bimodal weight distributions, generating insufficient agreement for a majority to endorse.

3. We also have shown that researchers can increase the level of agreement for a QOL index by weighting the components appropriately. Agreement is maximized by using the average weights from a survey of individuals' importances. Alternatively, if no surveys exist, equal weighting of indicators is the minimax estimator that minimizes disagreement even among diametrically opposed individuals. Note that in current practice, many QOL indices already use equal weighting of indicators, though their authors admit that they do not know whether this weighting is correct. The current results can now place current practice on a sound theoretical footing and show how it is possible to further increase agreement through surveys. (pp. 485-486)

The aggregation function we use for the index values for the eight indicators is a simple average or mean score. The simple average of any set of numbers is a familiar measure of the central tendency of the set, with familiar problems. Most notably, a mean (or average) score can provide a very misleading picture if one or a few figures in the set are wildly different from most others. Since the other standard measures of central tendency (medians and modes) can also provide misleading pictures and are somewhat less familiar than means, we use the latter. A fully developed CIW system would employ all three kinds of central tendency measures depending on which seem most useful for different purposes. As well, such a system would have a wide variety of distribution or variation measures. In this report, given our limited resources, the focus is only on national averages. In the future there should be reports explaining variation among regions, provinces, municipalities, and rural areas as well as among men and women, people with different incomes, ages, health status, levels of education, and so on.

Inspection of the eight distinct indicators reveals that the largest improvements occurred for the incidence of long term unemployment (160.4%) and the incidence of economic families in poverty (48.9%). The 86.1 figure in the last row of the first column means that there was a deterioration of 13.9% in the ratio of top to bottom quintile, after tax income of economic families (i.e., the gap between relatively rich and poor Canadians increased by 13.9%).

**Figure 1: Trends in Living Standards Indicators, 1994 to 2008**

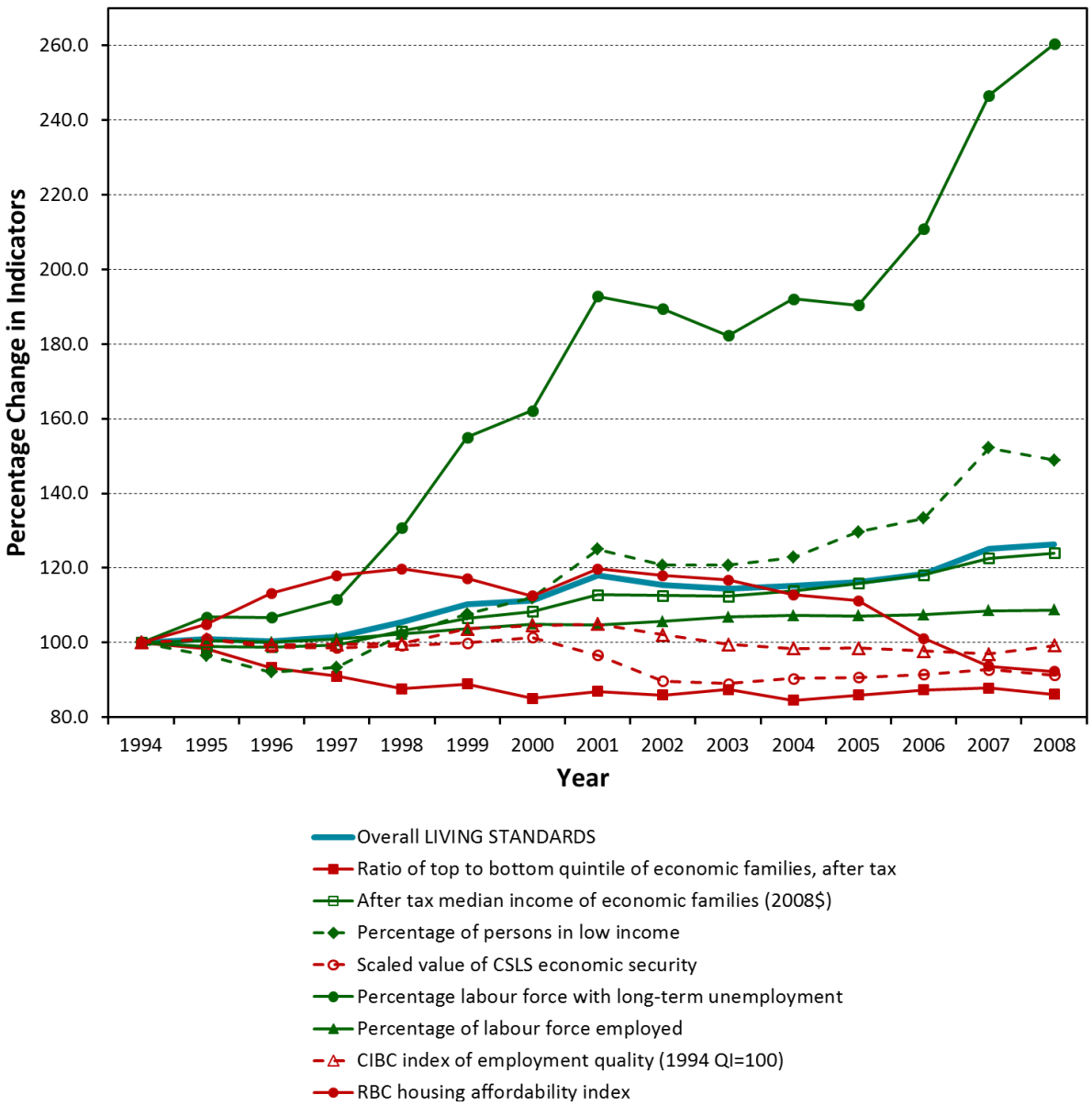



Figure 1 (above) illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW living standards domain plus the average of all of the eight indicators taken collectively (i.e., the CIW Domain score for living standards) across the 1994 to 2008 period. The figure clearly illustrates which headliners increased and which decreased the quality of life or wellbeing of Canadians with respect to the living standards domain. The top line far above the others represents the improvement in the incidence of long-term unemployment. The incidence of poverty is also above the CIW Domain score since 2001. The other trend lines are below the Domain trend beginning about 2004.

## *Healthy Populations*

The headliner indicators used for the *Healthy Populations Domain* (Table 2a) include three negative indicators and five positive indicators. They are listed here in the order in which they appear in the headings of the columns of Table 2a: (1p) percentage rating their own health as excellent or very good, (2n) percentage with diabetes, (3p) life expectancy at birth, (4n) teen smoking rate (aged 12 to 19 years, percentage daily or occasional smokers), (5n) percentage with probable depression, (6p) patient satisfaction with overall health services (percentage rating services as excellent or good), (7p) influenza immunization rate (aged 65 years and older, percentage saying "yes") and (8p) Average Health-Adjusted Life Expectancy (HALE) for population 15 years old and older (percentage of remaining years expected to be lived in good health).



**Table 2a: Trends in Healthy Populations Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Healthy Populations[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1p | 2n | 3p | 4n | 5n | 6p | 7p | 8p |
| **1994** | **63.1** | **3.0** | **78.2** | **20.9** | **5.3** | 84.4 | 47.9 | **85.3** |
| **1995** | 63.3 | 3.1 | 78.4 | 21.3 | 4.7 | 84.4 | 47.9 | 86.2 |
| **1996** | **63.4** | **3.2** | **78.6** | **21.6** | **4.1** | 84.4 | **47.9** | **87.1** |
| **1997** | 64.3 | 3.4 | 78.8 | 20.5 | 4.3 | 84.4 | 50.9 | 84.7 |
| **1998** | **65.2** | **3.5** | **79.0** | **19.4** | **4.5** | 84.4 | 53.9 | **82.2** |
| **1999** | 63.9 | 3.7 | 79.2 | 19.2 | 5.4 | 84.4 | 57.0 | 82.1 |
| **2000** | 62.7 | 3.9 | 79.4 | 18.9 | 6.3 | 84.4 | 60.0 | 81.9 |
| **2001** | **61.4** | **4.1** | **79.6** | **18.7** | **7.2** | **84.4** | **63.0** | **81.7** |
| **2002** | 59.9 | 4.4 | 79.8 | 16.8 | 6.6 | 85.6 | 62.7 | 82.8 |
| **2003** | **58.4** | **4.6** | **79.9** | **14.9** | **5.9** | **86.8** | **62.4** | **83.9** |
| **2004** | 59.3 | 4.8 | 80.2 | 13.5 | 5.6 | 86.0 | 64.5 | 82.9 |
| **2005** | **60.1** | **4.9** | **80.4** | **12.1** | **5.2** | **85.2** | **66.5** | **82.0** |
| **2006** | 59.9 | 5.4 | **80.8** | 12.1 | 5.6 | 86.0 | 65.4 | 82.0 |
| **2007** | **59.6** | **5.8** | 80.8 | **12.0** | **6.0** | **86.8** | **64.3** | 82.0 |
| **2008** | **58.9** | **5.9** | 80.8 | **11.4** | 6.0 | 86.8 | 64.3 | 82.0 |

[a] *Key:* 1p = Percentage self-rated health as excellent or very good
2n = Percentage with self-reported diabetes
3p = Life expectancy at birth, years
4n = Percentage of daily or occasional smokers among teens aged 12 to 19 years
5n = Percentage with probable depression
6p = Percentage rating patient health services as excellent or good
7p = Percentage aged 65 years or more getting influenza immunization
8p = Avg. number of remaining years expected to be lived in good health (avg. HALE 15+)

* Data which are *not* in bold were obtained by imputation. See text for Table 2a.

Figures entered in **boldface** were taken directly from surveys and/or administrative reports, while those in regular face *between* boldface entries are linear interpolations. When the first year with available real data was after 1994, the value of the first year with available data was used for previous years. When the last year of real data is before 2008, the most recent value of real data is simply repeated. These types of imputations are used in all of the following data tables. While Table 1a had 100% real data, Table 2a has 44%. Because the statistics are based on surveys taken roughly every other year and the numbers, which are aggregated population measures, do not change radically year by year, the 44% figure is not as troublesome as it may initially appear.



**Table 2b: Index of Healthy Populations Indicators for Canada, 1994 to 2008**

| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1p | 2n | 3p | 4n | 5n | 6p | 7p | 8p | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 100.2 | 96.8 | 100.3 | 98.4 | 112.8 | 100.0 | 100.0 | 101.1 | **101.2** |
| **1996** | 100.5 | 93.8 | 100.5 | 96.8 | 129.3 | 100.0 | 100.0 | 102.2 | **102.9** |
| **1997** | 101.9 | 89.6 | 100.8 | 102.0 | 123.3 | 100.0 | 106.3 | 99.3 | **102.9** |
| **1998** | 103.3 | 85.7 | 101.0 | 107.7 | 117.8 | 100.0 | 112.6 | 96.5 | **103.1** |
| **1999** | 101.3 | 81.1 | 101.3 | 109.0 | 98.1 | 100.0 | 118.9 | 96.3 | **100.8** |
| **2000** | 99.3 | 76.9 | 101.5 | 110.4 | 84.1 | 100.0 | 125.2 | 96.0 | **99.2** |
| **2001** | 97.3 | 73.2 | 101.8 | 111.8 | 73.6 | 100.0 | 131.5 | 95.8 | **98.1** |
| **2002** | 94.9 | 69.0 | 102.0 | 124.4 | 80.9 | 101.4 | 130.9 | 97.1 | **100.1** |
| **2003** | 92.6 | 65.2 | 102.2 | 140.3 | 89.8 | 102.8 | 130.3 | 98.4 | **102.7** |
| **2004** | 93.9 | 63.2 | 102.5 | 154.8 | 95.5 | 101.9 | 134.6 | 97.2 | **105.4** |
| **2005** | 95.2 | 61.2 | 102.8 | 172.7 | 101.9 | 100.9 | 138.8 | 96.1 | **108.7** |
| **2006** | 94.8 | 56.1 | 103.3 | 173.4 | 94.6 | 101.9 | 136.5 | 96.1 | **107.1** |
| **2007** | 94.5 | 51.7 | 103.3 | 174.2 | 88.3 | 102.8 | 134.2 | 96.1 | **105.7** |
| **2008** | 93.3 | 50.8 | 103.3 | 183.3 | 88.3 | 102.8 | 134.2 | 96.1 | **106.6** |

[a] *Key:* 1p = Percentage self-rated health as excellent or very good
2n = Percentage with self-reported diabetes
3p = Life expectancy at birth, years
4n = Percentage of daily or occasional smokers among teens aged 12 to 19 years
5n = Percentage with probable depression
6p =Percentage rating patient health services as excellent or good
7p = Percentage aged 65 years or more getting influenza immunization
8p = Avg. number of remaining years expected to be lived in good health (avg. HALE 15+)

[b] Average of healthy populations indicators

Inspection of the last line of the final column in Table 2b reveals that there was a 6.6% increase in the CIW health domain index in the period from 1994 to 2008. Inspection of the eight distinct indicators over the period reveals that there was an 83.3% improvement in the percentage of

teen-age smokers and a 34.2% improvement in the influenza immunization rate for older people. Examination of the negative indicators reveals a substantial increase of 49.2% in the self-reported prevalence of diabetes. Since the raw figures in Table 2a show that only 3.0% to 5.8% of the population reported having diabetes, the 49.2% increase in reported diabetes seems to have a greater impact on the CIW health domain index than its actual impact on the whole population. This apparent distortion is real and largely a consequence of our initial assumption to treat all headline indicators as equally important. So far, a good strategy for avoiding these sorts of distortions has not been found.



**Figure 2: Trends in Healthy Population Indicators, 1994 to 2008**

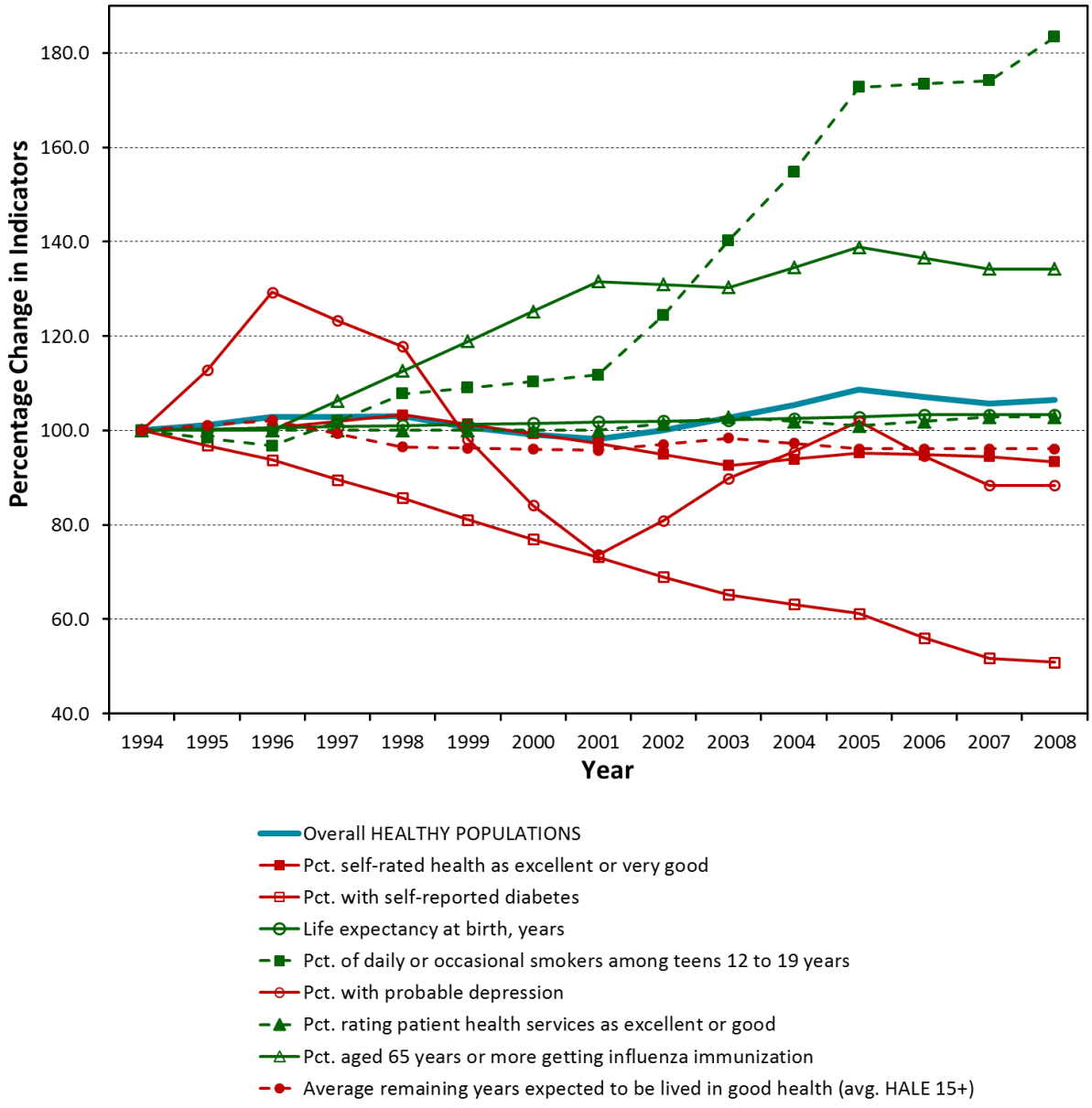


Figure 2 illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW healthy population domain plus the average of all the eight indicators taken collectively (i.e., the CIW Domain score for a healthy population) across the 1994 to 2008 period. The top line above the others represents the improvement in the teen smoking rate, followed by the trend in the influenza immunization rate for people 65 years or older. Beginning in 2004, the other six trend lines are below the CIW domain line. The diabetes prevalence trend line is below all others over the whole period, and increasingly so from 2001 forward.



## *Community Vitality*

Of 11 headliners recommended by Scott (2010), the eight used for the *Community Vitality Domain* composite and the CIW (Table 3a) include two negative indicators and six positive indicators. They are listed here in the order in which they appear in the headings of the columns of Table 3a: (1p) percentage reporting participation in organized group activities, (2p) percentage with six or more close friends, (3n) property crime rates per 100,000 population, (4n) violent crime rates per 100,000, (5p) percentage who feel safe walking alone after dark, (6p) percentage disagreeing that they worry less about the needs of others, (7p) percentage reporting they provide unpaid help to others on their own and (8p) percentage reporting a very or somewhat strong sense of belonging to their community. Thirty-eight per cent of the statistics in Table 3a represent real data.

A cursory review of the relatively few real data points for several columns reveals the difficulty encountered when one attempts to construct continuous time series of many things that matter to Canadians using on-the-shelf data, whether it is drawn from Statistics Canada or other sources. While we had an abundance of available time series with real data points for every headliner for every year from 1994 to 2008 in the living standards domain, our options shrunk to every other year in the healthy population domain and shrunk dramatically again in the community vitality domain. For one headliner in the latter domain (2p), we have real data for only three time periods. For five more headliners (1p, 5p to 8p), we have real data for four non-identical time periods. For (5p) a real figure from a survey in 1993 is used as our entry for 1994.

**Table 3a: Trends in Community Vitality Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Community Vitality[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1p | 2p[b] | 3n[c] | 4n | 5p | 6p | 7p | 8p |
| **1994** | 51.0 | 39.7 | 5,692 | 1,345 | **71.9** | **27.0** | 73.0 | 57.9 |
| **1995** | 51.0 | 39.7 | 5,692 | 1,345 | 72.3 | 31.7 | 73.0 | 57.9 |
| **1996** | 51.0 | **36.7** | 5,692 | 1,345 | 72.7 | 36.3 | 73.0 | 57.9 |
| **1997** | **51.0** | 36.3 | 5,692 | 1,345 | 73.0 | **41.0** | **73.0** | 57.9 |
| **1998** | 51.0 | 35.9 | **5,692** | **1,345** | 73.4 | 41.0 | 74.3 | 57.9 |
| **1999** | 51.0 | 35.5 | **5,345** | **1,440** | **73.8** | **41.0** | 75.7 | 57.9 |
| **2000** | **51.0** | 35.0 | **5,189** | **1,494** | 74.3 | 41.3 | **77.0** | 57.9 |
| **2001** | 54.3 | 34.6 | **5,124** | **1,473** | 74.9 | 41.5 | 78.5 | **57.9** |
| **2002** | 57.7 | 34.2 | **5,080** | **1,441** | 75.4 | 41.8 | 80.0 | 60.9 |
| **2003** | **61.0** | **33.8** | **5,299** | **1,435** | 76.0 | **42.0** | 81.5 | **63.9** |
| **2004** | 61.8 | 35.8 | **5,123** | **1,404** | **76.5** | 42.0 | **83.0** | 64.2 |
| **2005** | 62.6 | 37.8 | **4,884** | **1,389** | 77.2 | 42.0 | 83.3 | **64.4** |
| **2006** | 63.3 | 39.7 | **4,808** | **1,386** | 77.9 | 42.0 | 83.7 | 64.5 |
| **2007** | 64.1 | 41.7 | **4,519** | **1,352** | 78.6 | 42.0 | **84.0** | **64.6** |
| **2008** | **64.9** | **43.7** | **4,247** | **1,331** | **79.3** | 42.0 | 84.0 | **65.0** |

[a] *Key:* 1p = Percentage reporting participation in organized activities
2p = Percentage with 6 or more close friends
3n = Property crime rate per 100,000 population
4n = Violent crime rate per 100,000 population
5p = Percentage who feel safe walking alone after dark
6p = Percentage disagreeing that they worry less about the needs of others
7p = Percentage who provide unpaid help to others on their own
8p = Percentage reporting very or somewhat strong sense of belonging to community

* Data which are *not* in bold were obtained by imputation. See text for Table 2a.

Although the use of interpolations, imputations, and insertions of next-best figures in time series is not unusual for statistical analyses, it is unfortunate and undesirable, and the more one has of such things, the less valuable is one's monitoring system. All of us would much prefer to have real data to measure what matters most, and none of us enjoys having to choose between a very complete set of data revealing nothing of much importance to our wellbeing and a very incomplete set of data revealing something very important. While crime rates are certainly relevant to our wellbeing, as much cannot be said for some well-stocked time series in other domains. As emphasized in the CIW *Foundational Document*, one of the aims of our work is to assemble a set of indicators of wellbeing in our illustrative composite index that will persuade others with greater resources to make the investments required to ensure that those coming after us will not have to make such choices, or at least will have to make them less often.

**Table 3b: Index of Community Vitality Indicators for Canada, 1994 to 2008**

| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1p | 2p | 3n | 4n | 5p | 6p | 7p | 8p | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 100.0 | 100.0 | 100.0 | 100.0 | 100.5 | 117.3 | 100.0 | 100.0 | **102.2** |
| **1996** | 100.0 | 92.4 | 100.0 | 100.0 | 101.1 | 134.6 | 100.0 | 100.0 | **103.5** |
| **1997** | 100.0 | 91.4 | 100.0 | 100.0 | 101.6 | 151.9 | 100.0 | 100.0 | **105.6** |
| **1998** | 100.0 | 90.4 | 100.0 | 100.0 | 102.1 | 151.9 | 101.8 | 100.0 | **105.8** |
| **1999** | 100.0 | 89.3 | 106.5 | 93.4 | 102.6 | 151.9 | 103.7 | 100.0 | **105.9** |
| **2000** | 100.0 | 88.3 | 109.7 | 90.0 | 103.4 | 152.8 | 105.5 | 100.0 | **106.2** |
| **2001** | 106.5 | 87.2 | 111.1 | 91.3 | 104.1 | 153.7 | 107.5 | 100.0 | **107.7** |
| **2002** | 113.1 | 86.2 | 112.0 | 93.3 | 104.9 | 154.6 | 109.6 | 105.2 | **109.9** |
| **2003** | 119.6 | 85.1 | 107.4 | 93.7 | 105.6 | 155.6 | 111.6 | 110.4 | **111.1** |
| **2004** | 121.1 | 90.1 | 111.1 | 95.8 | 106.4 | 155.6 | 113.7 | 110.8 | **113.1** |
| **2005** | 122.7 | 95.1 | 116.5 | 96.8 | 107.4 | 155.6 | 114.2 | 111.2 | **114.9** |
| **2006** | 124.2 | 100.1 | 118.4 | 97.0 | 108.3 | 155.6 | 114.6 | 111.4 | **116.2** |
| **2007** | 125.7 | 105.1 | 126.0 | 99.5 | 109.3 | 155.6 | 115.1 | 111.6 | **118.5** |
| **2008** | 127.3 | 110.1 | 134.0 | 101.1 | 110.3 | 155.6 | 115.1 | 112.3 | **120.7** |

[a] *Key:* 1p = Percentage reporting participation in organized activities
2p = Percentage with six or more close friends
3n = Property crime rate per 100,000 population
4n = Violent crime rate per 100,000 population
5p = Percentage who feel safe walking alone after dark
6p = Percentage disagreeing that they worry less about the needs of others
7p = Percentage who provide unpaid help to others on their own
8p = Percentage reporting very or somewhat strong sense of belonging to community

[b] Average of healthy populations indicators

Inspection of the last row of the final column in Table 3b reveals that there was a 20.7% increase in the CIW community vitality domain index in the period from 1994 to 2008. Inspection of the eight distinct indicators over the period reveals that for the two negative indicators, there was a 34.0% and 1.1% improvement, respectively, in property and violent crime rates. The percentage increase (10.3%) in numbers of people feeling safe walking alone after dark represented an improvement in perceived wellbeing that was less than that indicated by the objective property crime rates and more than that indicated by the objective violent crime rates. A 55.6% increase in the numbers of Canadians denying that they are so hard pressed that they worry less about others suggests at least some improvement in how hard pressed they feel themselves to be. One would expect some upward bias (social desirability bias) in response to the survey question.

**Figure 3: Trends in Community Vitality Indicators, 1994 to 2008**

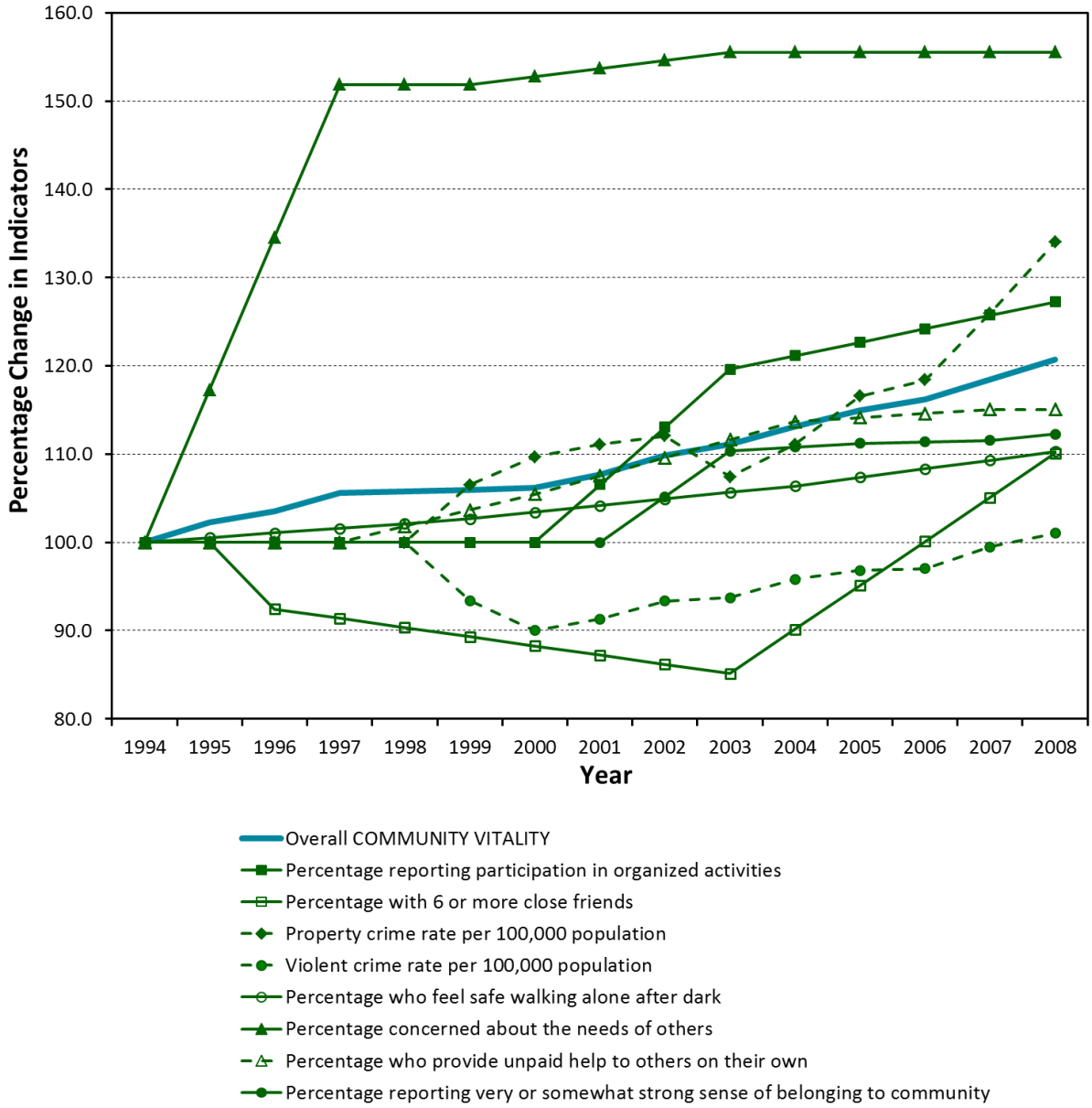



Figure 3 illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW community vitality domain plus the average of all the eight indicators taken collectively (i.e., the CIW domain score for community vitality) across the 1994 to 2008 period. The top line far above the others across the whole period represents the improvement in the percentage of people disagreeing that they worry less about the needs of others. The trend for people participating in organized group activities rises above all others beginning in 2002, but is overtaken by improvements in property crime rates by 2008. The trend for the percentage of people with six or more close friends is below all others from 1996 to 2005, and then rises above the improvements in violent crime rates. By 2008, five of the headliner scores are below the CIW Domain score.

## *Democratic Engagement*

Of 12 headliner indicators recommended by Moore *et al.* (2010), the eight used for the *Democratic Engagement Domain* composite and the CIW (Table 4a) include one negative and seven positive indicators. They are listed here in the order in which they appear in the headings of the columns of Table 4a: (1p) percentage of voter turnout at federal elections, (2n) percentage that are not interested at all in politics, (3p) percentage who strongly agree it is every citizen's duty to vote in federal elections, (4p) percentage reporting that they are very or fairly satisfied with the way democracy works in Canada, (5p) percentage reporting that policies of the federal government have made them better off, (6p) ratio of registered to eligible voters, (7p) percentage of women in Parliament and (8p) net Official Development Aid as a percentage of Gross National Income. Forty-four per cent of the figures in Table 4a represent real data.



**Table 4a: Trends in Democratic Engagement Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Democratic Engagement[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1p | 2n | 3p | 4p | 5p | 6p | 7p | 8p |
| **1994** | 67.0 | 9.7 | 75.0 | 57.2 | 6.1 | 0.90 | **18.0** | **0.43** |
| **1995** | 67.0 | 9.7 | 75.0 | 57.2 | 6.1 | 0.90 | **18.0** | **0.38** |
| **1996** | 67.0 | 9.7 | 75.0 | 57.2 | 6.1 | 0.90 | **18.0** | **0.32** |
| **1997** | **67.0** | **9.7** | 75.0 | **57.2** | **6.1** | **0.90** | **20.6** | **0.34** |
| **1998** | 65.1 | 9.6 | 75.0 | 59.0 | 10.8 | 0.93 | **20.6** | **0.30** |
| **1999** | 63.2 | 9.6 | 75.0 | 60.8 | 15.5 | 0.95 | **20.6** | **0.28** |
| **2000** | **61.3** | **9.5** | 75.0 | **62.6** | **20.2** | **0.98** | **20.5** | **0.25** |
| **2001** | 61.2 | 10.1 | 75.0 | 60.5 | 17.7 | 0.98 | **20.5** | **0.22** |
| **2002** | 61.1 | 10.6 | 75.0 | 58.4 | 15.2 | 0.98 | **20.5** | **0.28** |
| **2003** | 61.0 | 11.2 | 75.0 | 56.3 | 12.6 | 0.98 | **20.5** | **0.24** |
| **2004** | **60.9** | **11.7** | **75.0** | **54.2** | **10.1** | **0.98** | **21.1** | **0.27** |
| **2005** | 62.8 | 9.4 | 80.5 | 56.6 | 11.4 | 0.96 | **21.1** | **0.34** |
| **2006** | **64.7** | **7.1** | **86.0** | **59.0** | **12.6** | **0.93** | **20.7** | **0.29** |
| **2007** | 61.9 | 7.1 | 86.0 | 59.0 | 12.6 | 0.95 | **20.7** | **0.28** |
| **2008** | **59.1** | 7.1 | 86.0 | 59.0 | 12.6 | **0.96** | **22.4** | **0.32** |

[a] *Key:* 1p = Percentage of voter turnout at federal elections
2n = Percentage that are not interested in politics at all
3p = Percentage strongly agree it is every citizen's duty to vote in federal Elections
4p = Pct. reporting they are very/fairly satisfied with the way democracy works in Canada
5p = Pct. reporting that policies of the federal government have made them better off
6p = Ratio of registered to eligible voters
7p = Percentage of women in Parliament
8p = Net official development aid as a percentage of gross national income

* Data which are *not* in bold were obtained by imputation. See text for Table 2a.

Again, a review of the relatively few real data points for columns (2n) to (5p) reveals the difficulty encountered when one attempts to construct continuous time series of many things that matter to Canadians using on-the-shelf data. For one headliner (3p), we have real data for only two time periods and for three (2n, 4p, and 5p), we have real data for four time periods.

**Table 4b: Index of Democratic Engagement Indicators for Canada, 1994 to 2008**

| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1p | 2n | 3p | 4p | 5p | 6p | 7p | 8p | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 88.4 | **98.5** |
| **1996** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 74.4 | **96.8** |
| **1997** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 114.4 | 79.1 | **99.2** |
| **1998** | 97.2 | 100.7 | 100.0 | 103.1 | 177.0 | 103.0 | 114.4 | 69.8 | **108.2** |
| **1999** | 94.3 | 101.4 | 100.0 | 106.3 | 254.1 | 105.9 | 114.4 | 65.1 | **117.7** |
| **2000** | 91.5 | 102.1 | 100.0 | 109.4 | 331.1 | 108.9 | 113.9 | 58.1 | **126.9** |
| **2001** | 91.3 | 96.5 | 100.0 | 105.8 | 289.8 | 108.9 | 113.9 | 51.2 | **119.7** |
| **2002** | 91.2 | 91.5 | 100.0 | 102.1 | 248.4 | 108.9 | 113.9 | 65.1 | **115.1** |
| **2003** | 91.0 | 87.0 | 100.0 | 98.4 | 207.0 | 108.9 | 113.9 | 55.8 | **107.8** |
| **2004** | 90.9 | 82.9 | 100.0 | 94.8 | 165.6 | 108.9 | 117.2 | 62.8 | **102.9** |
| **2005** | 93.7 | 103.2 | 107.3 | 99.0 | 186.1 | 106.1 | 117.2 | 79.1 | **111.5** |
| **2006** | 96.6 | 136.6 | 114.7 | 103.1 | 206.6 | 103.3 | 115.0 | 67.4 | **117.9** |
| **2007** | 92.4 | 136.6 | 114.7 | 103.1 | 206.6 | 105.0 | 115.0 | 65.1 | **117.3** |
| **2008** | 88.2 | 136.6 | 114.7 | 103.1 | 206.6 | 106.7 | 124.4 | 74.4 | **119.3** |

[a] *Key:* 1p = Percentage of voter turnout at federal elections
2n = Percentage that are not interested in politics at all
3p = Percentage strongly agree it is every citizen's duty to vote in federal Elections
4p = Pct. reporting they are very/fairly satisfied with the way democracy works in Canada
5p = Pct. reporting that policies of the federal government have made them better off
6p = Ratio of registered to eligible voters
7p = Percentage of women in Parliament
8p = Net official development aid as a percentage of gross national income

[b] Average of democratic engagement indicators

Inspection of the last row of the final column in Table 4b reveals that there was a 19.3% increase in the CIW democratic engagement domain index in the period from 1994 to 2008. Inspection of the eight distinct indicators over the period reveals that for the one negative indicator, there was a 36.6% improvement in the percentage of sampled Canadians who reported that they were not at all interested in politics. A single indicator with only four real data values had the greatest impact on the overall average domain score. The percentage of sampled Canadians reporting that policies of the federal government made them better off began at 6.1% in 1997, rose dramatically to 20.2% in 2000, and then fell back to 12.6% in 2006. We do not have a good explanation for these relatively low levels of appreciation for federal

government activities or for the great variation of scores over the whole period. In the light of reviews like those of Cummins (1996), relatively low results were not unexpected, but we had no reason to expect the wide variation.

**Figure 4: Trends in Democratic Engagement Indicators, 1994 to 2008**

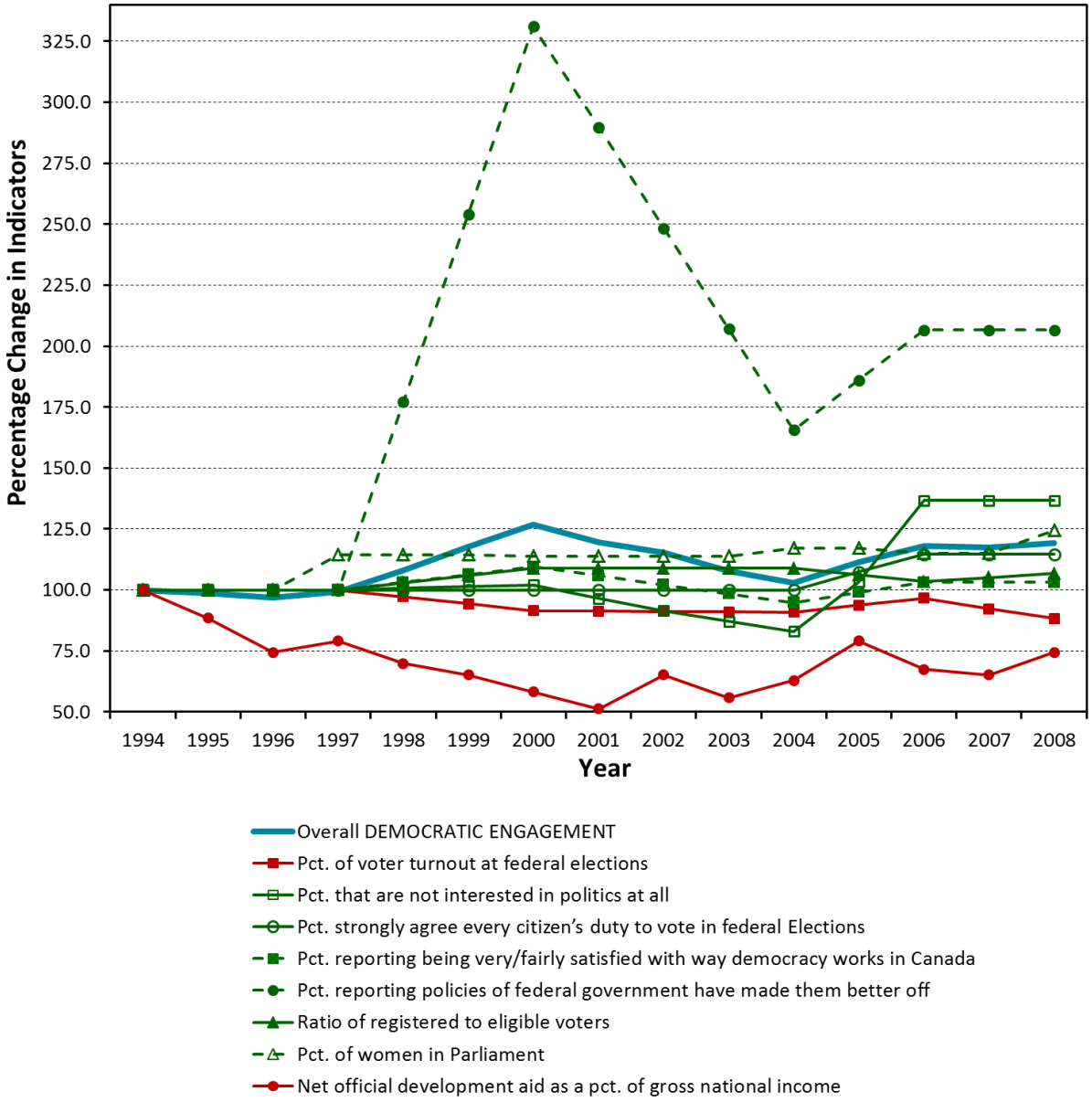



Figure 4 illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW democratic engagement domain plus the CIW Domain score across the 1994 to 2008 period. The spiked top line far above the others across the period from 1997 represents the improvement in the percentage of people reporting that the policies of the federal government have made them better off. The bottom line across the whole period represents the deterioration of Canada's Official Development Aid.

## *Leisure and Culture*

Table 5a contains data for the eight headline indicators recommended by Smale *et al.* (2010) in the *Leisure and Culture Domain* for the 1994 to 2008 period. All eight are positive indicators. They are listed here in the order in which they appear in the headings of the columns of Table 5a: (1p) average percentage of time spent on the previous day in social leisure activities, (2p) average percentage of time spent on the previous day in arts and culture activities, (3p) average number of hours in the past year volunteering for culture and recreation organizations, (4p) average monthly frequency of participation in physical activity lasting over 15 minutes, (5p) average attendance per performance in the past year at all performing arts performances, (6p) average visitations per site in the past year to all National Parks and National Historic Sites, (7p) average number of nights away per trip in the past year on vacation trips to destinations over 80 km from home and (8p) expenditures in the past year on all aspects of culture and recreation as a percentage of total household expenditures.



**Table 5a: Trends in Leisure and Culture Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Leisure and Culture[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **1p** | **2p** | **3p** | **4p** | **5p** | **6p** | **7p** | **8p** |
| **1994** | **15.23** | **4.47** | 46.17 | **20.74** | 338.34 | 219,773 | 2.65 | 20.46 |
| **1995** | 15.18 | 4.45 | 46.17 | 21.41 | 338.34 | **219,773** | 2.65 | 20.46 |
| **1996** | 15.13 | 4.44 | 46.17 | **22.08** | 338.34 | **214,681** | **2.65** | 20.46 |
| **1997** | 15.08 | 4.42 | **46.17** | 22.34 | 338.34 | **202,091** | **2.76** | **20.46** |
| **1998** | **15.03** | **4.40** | 45.94 | 22.59 | 338.34 | **205,569** | **2.63** | **20.40** |
| **1999** | 14.66 | 4.38 | 45.70 | **22.85** | 338.34 | **211,355** | **2.56** | 20.44 |
| **2000** | 14.28 | 4.35 | **45.47** | 22.98 | 338.34 | **219,672** | **2.71** | **20.47** |
| **2001** | 13.91 | 4.33 | 44.52 | **23.11** | **338.34** | **183,064** | **2.67** | **22.49** |
| **2002** | 13.53 | 4.30 | 43.56 | **24.92** | **324.41** | **186,583** | **2.52** | **21.76** |
| **2003** | 13.16 | 4.28 | 42.61 | **25.98** | 320.54 | **191,685** | **2.47** | **22.01** |
| **2004** | 12.78 | 4.25 | **41.65** | **22.44** | **316.66** | **176,584** | **2.58** | **21.63** |
| **2005** | **12.41** | **4.23** | 40.53 | **25.83** | 328.43 | **168,798** | 2.68 | **21.91** |
| **2006** | 12.41 | 4.23 | 39.40 | 25.83 | **340.19** | **174,355** | **2.78** | **21.32** |
| **2007** | 12.41 | 4.23 | 38.28 | 25.83 | 340.19 | **171,539** | **2.95** | 21.32 |
| **2008** | 12.41 | 4.23 | **37.15** | 25.83 | 340.19 | **172,678** | 2.95 | 21.32 |

[a] *Key:* 1p = Average percentage of time spent on the previous day in social leisure activities
2p = Average percentage of time spent on the previous day in arts and culture activities
3p = Avg. number of hours in past year volunteering for culture and recreation orgs.
4p = Avg. monthly frequency of participation in physical activity lasting over 15 minutes
5p = Avg. attendance per performance in past year at all performing arts performances
6p = Average visitation per site in past year to all National Parks and National Historic Sites
7p = Avg. nights away/trip in past year on vacations to destinations over 80 km from home
8p = Expenditures in past year on all culture/recreation as pct. of total household expend.

* Data which are *not* in bold were obtained by imputation. See text for Table 2a.

Forty-seven per cent of the figures in Table 5a represent real data. The headliner with the greatest percentage of real data points is that for the average visitations per site in the past year to all National Parks and National Historic Sites. Each of the first two columns has only three real data points, and the third column has four.

**Table 5b: Index of Leisure and Culture Indicators for Canada, 1994 to 2008**

| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1p | 2p | 3p | 4p | 5p | 6p | 7p | 8p | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 99.7 | 99.6 | 100.0 | 103.2 | 100.0 | 100.0 | 100.0 | 100.0 | **100.3** |
| **1996** | 99.3 | 99.2 | 100.0 | 106.5 | 100.0 | 97.7 | 100.0 | 100.0 | **100.3** |
| **1997** | 99.0 | 98.8 | 100.0 | 107.7 | 100.0 | 92.0 | 104.2 | 100.0 | **100.2** |
| **1998** | 98.7 | 98.4 | 99.5 | 108.9 | 100.0 | 93.5 | 99.2 | 99.7 | **99.8** |
| **1999** | 96.2 | 97.9 | 99.0 | 110.2 | 100.0 | 96.2 | 96.6 | 99.9 | **99.5** |
| **2000** | 93.8 | 97.3 | 98.5 | 110.8 | 100.0 | 100.0 | 102.3 | 100.0 | **100.3** |
| **2001** | 91.3 | 96.8 | 96.4 | 111.4 | 100.0 | 83.3 | 100.8 | 109.9 | **98.7** |
| **2002** | 88.9 | 96.3 | 94.3 | 120.2 | 95.9 | 84.9 | 95.1 | 106.4 | **97.7** |
| **2003** | 86.4 | 95.7 | 92.3 | 125.3 | 94.7 | 87.2 | 93.2 | 107.6 | **97.8** |
| **2004** | 83.9 | 95.2 | 90.2 | 108.2 | 93.6 | 80.3 | 97.4 | 105.7 | **94.3** |
| **2005** | 81.5 | 94.6 | 87.8 | 124.5 | 97.1 | 76.8 | 101.1 | 107.1 | **96.3** |
| **2006** | 81.5 | 94.6 | 85.3 | 124.5 | 100.5 | 79.3 | 104.9 | 104.2 | **96.9** |
| **2007** | 81.5 | 94.6 | 82.9 | 124.5 | 100.5 | 78.1 | 111.3 | 104.2 | **97.2** |
| **2008** | 81.5 | 94.6 | 80.5 | 124.5 | 100.5 | 78.6 | 111.3 | 104.2 | **97.0** |

[a] *Key:* 1p = Average percentage of time spent on the previous day in social leisure activities
2p = Average percentage of time spent on the previous day in arts and culture activities
3p = Avg. number of hours in past year volunteering for culture and recreation orgs.
4p = Avg. monthly frequency of participation in physical activity lasting over 15 minutes
5p = Avg. attendance per performance in past year at all performing arts performances
6p = Average visitation per site in past year to all National Parks and National Historic Sites
7p = Avg. nights away/trip in past year on vacations to destinations over 80 km from home
8p = Expenditures in past year on all culture/recreation as pct. of total household expend.

[b] Average of democratic engagement indicators

The last figure in the final column of Table 5b reveals a small decline of 3.0% in the average value of leisure and culture headline indicators (i.e., in the CIW leisure and culture domain index in the period from 1994 to 2008). Inspection of the last row of the eight distinct indicators in this table reveals that greatest losses occurred for average visitations per site in the past year to all National Parks and National Historic sites (21.4%), followed by the average number of hours in the past year volunteering for culture and recreation organizations (19.5%). Since the number of parks and sites increased over the whole period, some of the decrease in

the average number of visitations to such places is in part the result of this increase rather than the result of relatively fewer people making such visits, although in absolute terms, visits have declined as well. Greatest gains were made for the average monthly frequency of participation in physical activity lasting over 15 minutes (24.5%), followed by the average number of nights away per trip in the past year on vacation trips to destinations over 80 km from home (11.3%).

**Figure 5: Trends in Leisure and Culture Indicators, 1994 to 2008**

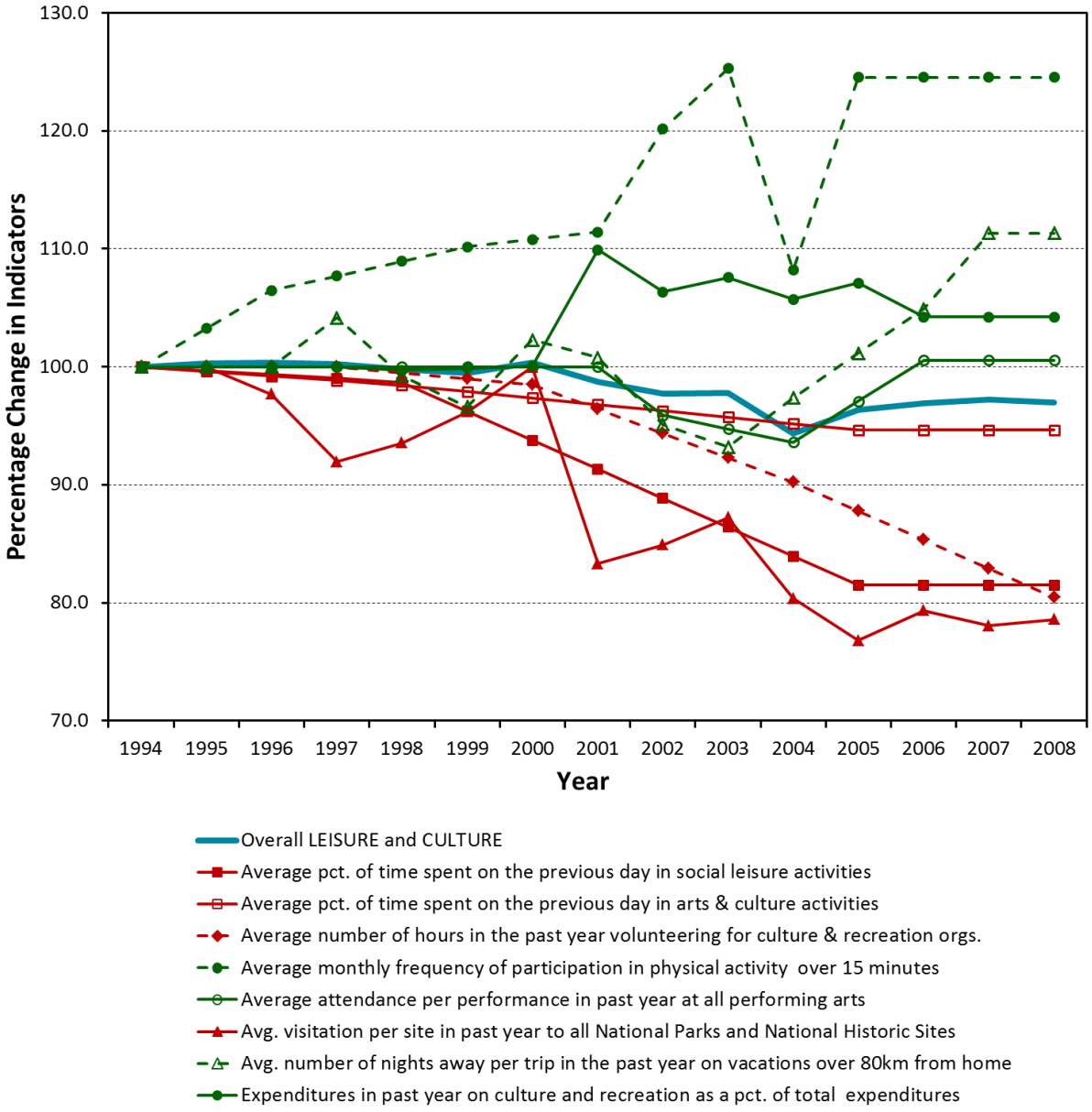



Figure 5 (above) illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW leisure and culture domain plus the CIW domain score for leisure and culture across the 1994 to 2008 period. The somewhat variable line above all others over

the whole period represents the average monthly frequency of participation in physical activity lasting over 15 minutes. The somewhat variable line below all others most of the period represents the average visitation per site in the past year to all National Parks and National Historic Sites. The line representing the average percentage of time spent on the previous day in social leisure activities shows almost uniform deterioration annually from 1998 forward.



## *Time Use*

Table 6a contains data for eight of nine headline indicators recommended by Brooker and Hyman (2010) in the *Time Use Domain* for the 1994 to 2008 period. Four are positive and four are negative indicators. They are listed here in the order in which they appear in the headings of the columns of Table 6a: (1n) percentage of people 20 to 64 years old working 50 hours or more per week, (2n) percentage of people 20 to 64 years old reporting high levels of time pressure, (3n) percentage of people 20 to 64 years old giving unpaid care to seniors, (4p) percentage of people 65 years old and older reporting daily active leisure activities, (5p) percentage of people 65 years old and older reporting annual formal volunteering activities, (6n) percentage of people 12 to 17 years old spending 2 hours or more per day on TV or video games, (7p) percentage of people 6 to 9 years old participating in structured activities at least once a week and (8p) percentage of people 3 to 5 years old being read to daily by parents.

Only 33% of the figures in Table 6a represent real data, which implies that more imputations were required in this domain than in any other. The headliner with the greatest percentage of real data points is that for the percentage of people 20 to 64 years old working 50 hours or more per week. Two indicators have only two data points (i.e., percentage of people 65 years old and older reporting annual formal volunteering activities and percentage of people 12 to 17 years old spending two hours or more per day on TV or video games) and three more have only three data points (i.e., percentage of people 20 to 64 years old reporting high levels of time pressure, percentage of people 20 to 64 years old giving unpaid care to seniors, and percentage of people 65 years old and older reporting daily active leisure activities).

**Table 6a: Trends in Time Use Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Time Use[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1n | 2n | 3n | 4p | 5p | 6n | 7p | 8p |
| **1994** | **14.7** | **16.4** | 17.4 | **89.7** | 31.6 | 27.2 | 75.7 | **61.7** |
| **1995** | **14.3** | 17.8 | 17.4 | 89.7 | 31.6 | 27.2 | 75.7 | 62.1 |
| **1996** | **14.9** | 19.2 | **17.4** | 89.6 | 31.6 | 27.2 | **75.7** | **62.4** |
| **1997** | **14.8** | 20.5 | 17.8 | 89.6 | 31.6 | 27.2 | 76.2 | 64.3 |
| **1998** | **14.4** | **21.9** | 18.2 | **89.5** | 31.6 | 27.2 | **76.6** | **66.2** |
| **1999** | **14.4** | 21.6 | 18.5 | 89.0 | 31.6 | 27.2 | 76.9 | 64.5 |
| **2000** | **14.4** | 21.2 | 18.9 | 88.4 | 31.6 | 27.2 | **77.1** | **62.8** |
| **2001** | **13.3** | 20.9 | **19.3** | 87.9 | 31.6 | 27.2 | 78.0 | 63.2 |
| **2002** | **13.0** | 20.6 | 19.3 | 87.3 | 31.6 | 27.2 | 78.9 | **63.5** |
| **2003** | **12.2** | 20.3 | 19.4 | 86.8 | 31.6 | **27.2** | 79.8 | 62.0 |
| **2004** | **12.7** | 19.9 | 19.4 | 86.2 | **31.6** | 28.3 | 80.7 | **60.4** |
| **2005** | **13.0** | **19.6** | 19.5 | **85.7** | 32.6 | 29.5 | 81.6 | 61.5 |
| **2006** | **12.5** | 19.6 | **19.5** | 85.7 | 33.5 | 30.6 | **82.5** | **62.6** |
| **2007** | **12.8** | 19.6 | 19.5 | 85.7 | **34.5** | **31.7** | 82.5 | 62.6 |
| **2008** | **12.1** | 19.6 | 19.5 | 85.7 | 34.5 | 31.7 | 82.5 | 62.6 |

[a] *Key:* 1n = Percentage of 20 to 64 year olds working over 50 hours per week
2n = Percentage of 20 to 64 year olds reporting high levels of time pressure
3n = Percentage of 20 to 64 years old giving unpaid care to seniors
4p = Percentage of 65 years and older reporting daily active leisure activities
5p = Percentage of 65 years and older reporting annual formal volunteering activities
6n =Pct. of 12 to 17 year olds spending 2 hours or more per day on TV or video games
7p = Percentage of 6 to 9 year olds having weekly or more structured activities
8p = Percentage of 3 to 5 year olds read to daily by parents

* Data which are *not* in bold were obtained by imputation. See text for Table 2a.

The last figure in the final column of Table 6b reveals a tiny decrease of 0.6% in the average value of the eight time use headline indicators (i.e., in the CIW time use domain index in the period from 1994 to 2008). In fact, across the whole period, there is very little change in the average figures. However, inspection of the last row of the eight distinct indicators in this table reveals some substantial changes. The greatest improvements occurred for the percentage of people 20 to 64 years old working 50 hours or more per week (21.5%), followed by a virtual tie between the percentage of people 65 years old and older reporting annual formal volunteering activities (9.2%) and the percentage of 6 to 9 years olds participating in structured activities at least once a week (9.0%). Greatest losses occurred for the percentage of people 20 to 64 years old reporting high levels of time pressure (16.3%), followed by the percentage of people 12 to 17 years old spending two hours or more per day on TV or video games (14.2%).

**Table 6b: Index of Time Use Indicators for Canada, 1994 to 2008**

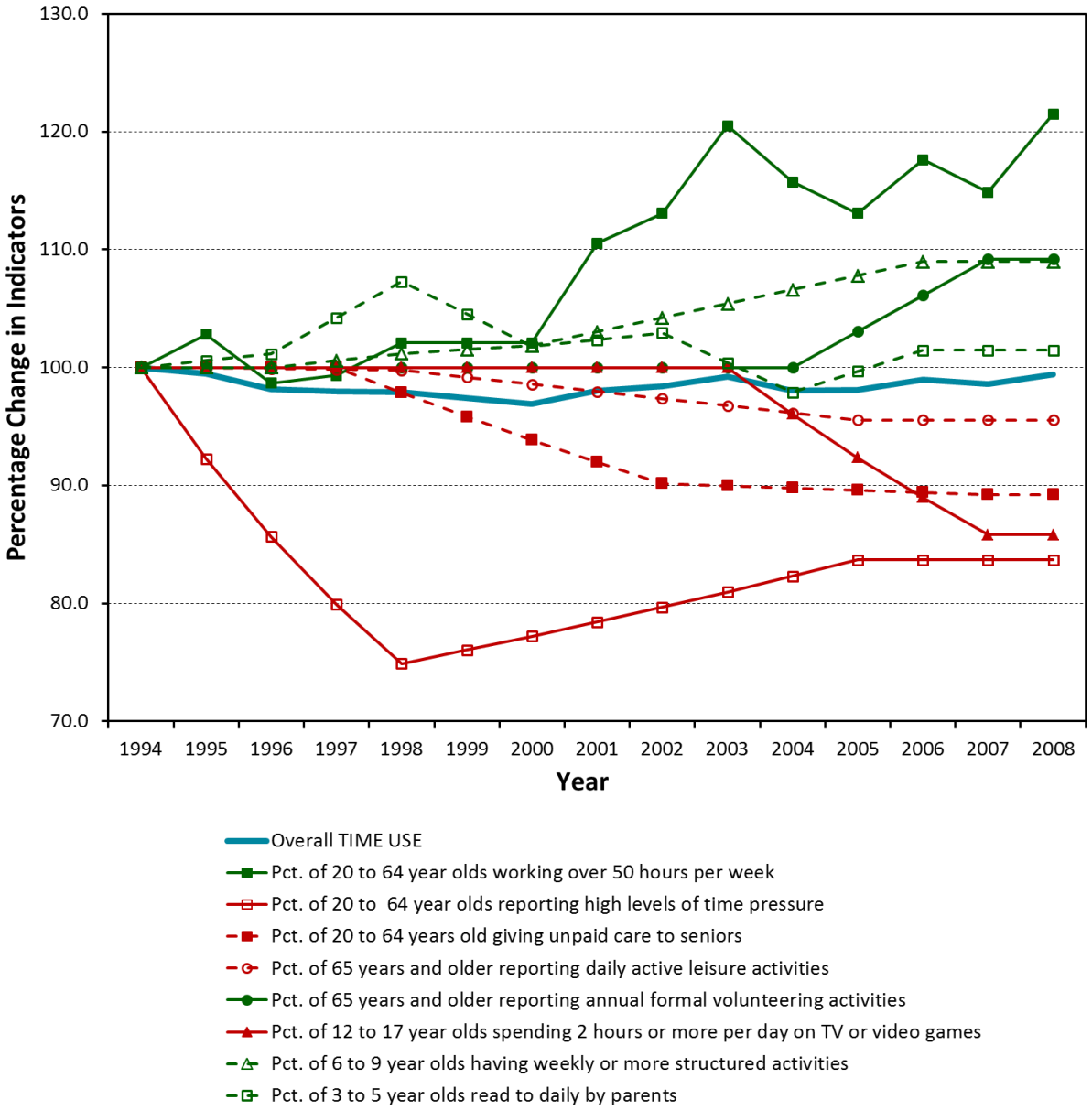
| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1n | 2n | 3n | 4p | 5p | 6n | 7p | 8p | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 102.8 | 92.3 | 100.0 | 99.9 | 100.0 | 100.0 | 100.0 | 100.6 | **99.4** |
| **1996** | 98.7 | 85.6 | 100.0 | 99.9 | 100.0 | 100.0 | 100.0 | 101.1 | **98.2** |
| **1997** | 99.3 | 79.9 | 100.0 | 99.8 | 100.0 | 100.0 | 100.6 | 104.2 | **98.0** |
| **1998** | 102.1 | 74.9 | 97.9 | 99.8 | 100.0 | 100.0 | 101.2 | 107.3 | **97.9** |
| **1999** | 102.1 | 76.0 | 95.8 | 99.2 | 100.0 | 100.0 | 101.5 | 104.5 | **97.4** |
| **2000** | 102.1 | 77.2 | 93.9 | 98.6 | 100.0 | 100.0 | 101.8 | 101.8 | **96.9** |
| **2001** | 110.5 | 78.4 | 92.0 | 98.0 | 100.0 | 100.0 | 103.0 | 102.4 | **98.0** |
| **2002** | 113.1 | 79.7 | 90.2 | 97.4 | 100.0 | 100.0 | 104.2 | 102.9 | **98.4** |
| **2003** | 120.5 | 81.0 | 90.0 | 96.8 | 100.0 | 100.0 | 105.4 | 100.4 | **99.2** |
| **2004** | 115.7 | 82.3 | 89.8 | 96.1 | 100.0 | 96.0 | 106.6 | 97.9 | **98.1** |
| **2005** | 113.1 | 83.7 | 89.6 | 95.5 | 103.1 | 92.4 | 107.8 | 99.7 | **98.1** |
| **2006** | 117.6 | 83.7 | 89.4 | 95.5 | 106.1 | 89.0 | 109.0 | 101.5 | **99.0** |
| **2007** | 114.8 | 83.7 | 89.2 | 95.5 | 109.2 | 85.8 | 109.0 | 101.5 | **98.6** |
| **2008** | 121.5 | 83.7 | 89.2 | 95.5 | 109.2 | 85.8 | 109.0 | 101.5 | **99.4** |

[a] *Key:* 1n = Percentage of 20 to 64 year olds working over 50 hours per week
2n = Percentage of 20 to 64 year olds reporting high levels of time pressure
3n = Percentage of 20 to 64 years old giving unpaid care to seniors
4p = Percentage of 65 years and older reporting daily active leisure activities
5p = Percentage of 65 years and older reporting annual formal volunteering activities
6n =Pct. of 12 to 17 year olds spending 2 hours or more per day on TV or video games
7p = Percentage of 6 to 9 year olds having weekly or more structured activities
8p = Percentage of 3 to 5 year olds read to daily by parents

[b] Average of democratic engagement indicators

Figure 6 illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW time use domain plus the CIW Domain score for time use across the 1994 to 2008 period. The line above all others most of the time beginning in 1996 represents improvements in the percentage of people aged 20 to 64 years working 50 or more hours per week. The lowest line represents significant deterioration in the percentage of people in this same age group reporting high levels of time pressure. Since 1997, this group also suffered some deterioration in the quality of their lives as a result of giving unpaid care to seniors.

**Figure 6: Trends in Time Use Indicators, 1994 to 2008**




## *Education*

Table 7a contains data for eight headline indicators recommended by Guhn, Gadermann, and Zumbo (2010) in the *Education Domain* for the 1994 to 2008 period. Six are positive and two are negative indicators. They are listed here in the order in which they appear in the headings of the columns of Table 7a: (1p) ratio of childcare spaces to children aged 0 to 5 years of age (%), (2p) percentage of children doing well on five developmental domains, at age 5 years, (3n)

ratio of students to educators in public schools, (4p) average of five social and emotional competence scores for 12 to 13 year olds, (5p) basic knowledge and skills index for 13 to 15 year olds, (6n) percentage of variation in PISA scores explained by socio-economic background, (7p) percentage of the 20 to 24 year old population completing high school, and (8p) percentage of the 25 to 64 year old population with a university degree.

**Table 7a: Trends in Education Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Education[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1p | 2p | 3n | 4p | 5p[b] | 6n | 7p | 8p |
| **1994** | 12.0 | **83.0** | 15.9 | 3.25 | 523.0 | 11.0 | **86.0** | **19.0** |
| **1995** | **12.0** | 83.0 | 15.9 | 3.25 | **523.0** | 11.0 | **87.0** | **19.0** |
| **1996** | 12.3 | **83.0** | 15.9 | **3.25** | 525.5 | 11.0 | **87.0** | **20.0** |
| **1997** | 12.7 | 84.0 | **15.9** | 3.23 | 528.0 | 11.0 | **88.0** | **21.0** |
| **1998** | **13.0** | **85.0** | **16.5** | **3.20** | 530.5 | 11.0 | **88.0** | **21.0** |
| **1999** | 13.7 | 85.5 | **16.2** | 3.19 | **533.0** | 11.0 | **89.0** | **22.0** |
| **2000** | 14.3 | **86.0** | **16.2** | **3.18** | 532.5 | **11.0** | **89.0** | **23.0** |
| **2001** | **15.0** | 86.5 | **15.9** | 3.17 | 532.0 | 11.0 | **89.0** | **24.0** |
| **2002** | 15.7 | **87.0** | **15.9** | **3.15** | 531.5 | 11.0 | **89.0** | **24.0** |
| **2003** | 16.3 | 86.5 | **15.9** | 3.14 | **531.0** | **11.0** | **89.0** | **25.0** |
| **2004** | **17.0** | **86.0** | **15.8** | **3.13** | 528.0 | 10.3 | **90.0** | **25.0** |
| **2005** | 18.0 | 86.0 | **15.5** | 3.13 | 525.0 | 9.7 | **91.0** | **26.0** |
| **2006** | **19.0** | **86.0** | **15.2** | **3.13** | **522.0** | **9.0** | **91.0** | **27.0** |
| **2007** | 19.5 | 86.0 | **14.7** | 3.13 | 522.0 | 9.0 | **91.0** | **27.0** |
| **2008** | **20.0** | 86.0 | 14.7 | 3.13 | 522.0 | 9.0 | 91.0 | **28.0** |

[a] *Key:* 1p = Ratio of childcare spaces to children aged 0 to 5 years of age
2p = Percentage of children doing well on five developmental domains, at age 5 years
3n = Ratio of students to educators in public schools
4p = Average of 5 social and emotional competence scores for 12 to 13 year olds
5p = Basic knowledge and skills index for 13 to 15 year olds
6n = Percentage of PISA scores explained by socio-economic background
7p = Percentage of 20 to 24 year olds in population completing high school
8p = Percentage of 25 to 64 year olds in population with a university degree

[b] mean of the PISA literacy score and the TIMSS math and science scores used for 1995 and 1999 were taken from 2000 survey

* Data which are *not* in bold were obtained by imputation. See text for Table 2a.

Fifty-five per cent of the figures in Table 7a represent real data. The headliner with the greatest percentage of real data points (all 15) is that for the percentage of people 25 to 64 years of age with a university degree. One indicator has only three data points (i.e., percentage of variation in PISA scores explained by socio-economic background) and one has four (i.e., basic knowledge and skills index for 13 to 15 year olds). All others have six or more real data points.

The last figure in the final column of Table 7b reveals an increase of 18.7% in the average value of the eight education headline indicators (i.e., in the CIW education domain index for the period from 1994 to 2008). Inspection of the last row of the eight distinct indicators in this table reveals some substantial changes. The greatest improvements occurred for the ratio of childcare spaces to children aged 0 to 5 years (66.7%), followed by the percentage of the 25 to 64 year old population with a university degree (47.4%). Only two headliners indicated losses over the whole period and one of those was practically negligible. Losses were indicated for the average of five social and emotional competence scores for 12 to 13 year olds (3.7%) and the basic knowledge and skills index (0.2%). On the whole, the education domain faired relatively well compared to all others.



**Table 7b: Index of Education Indicators for Canada, 1994 to 2008**

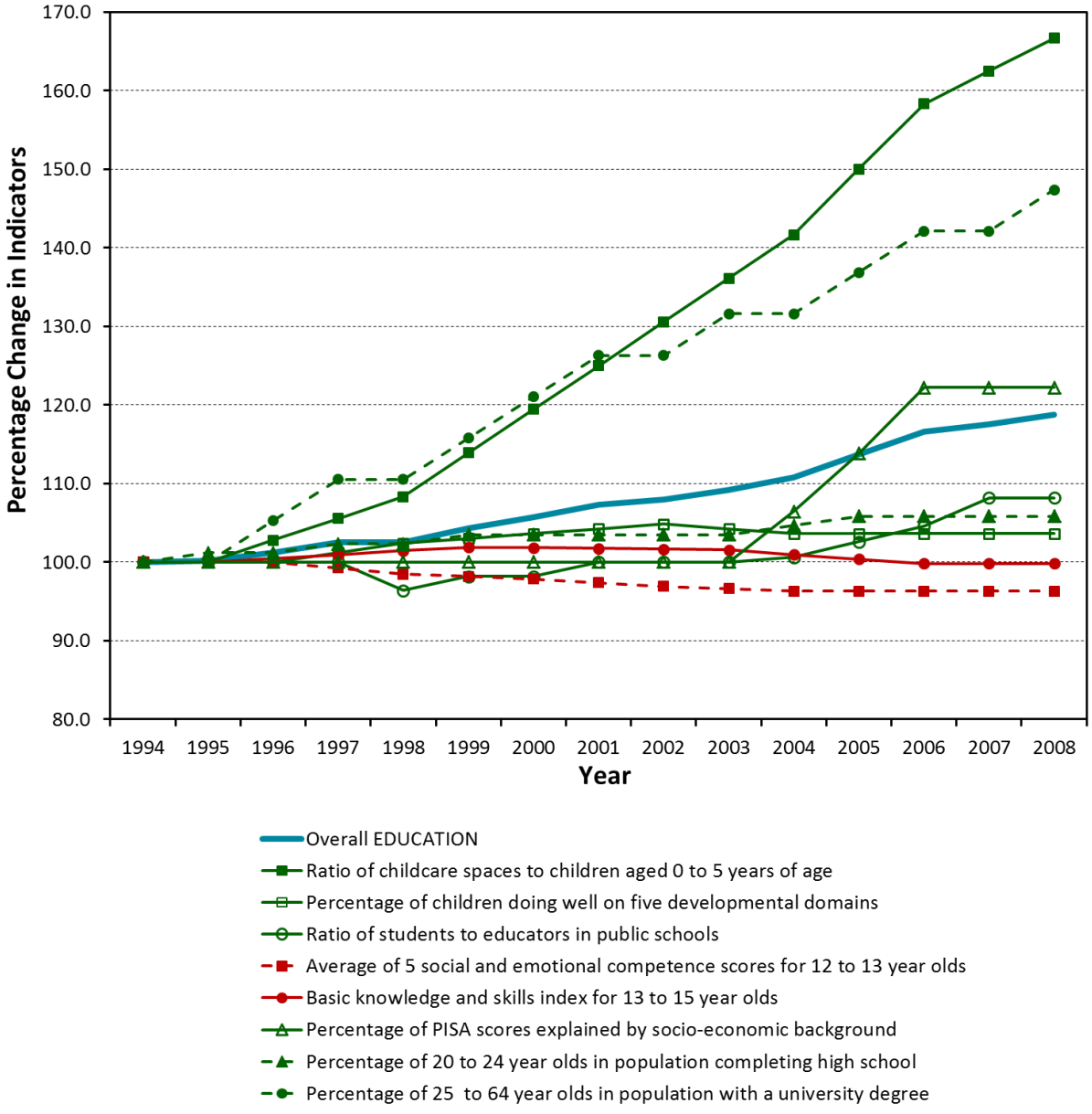
| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1p | 2p | 3n | 4p | 5p | 6n | 7p | 8p | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 101.2 | 100.0 | **100.1** |
| **1996** | 102.8 | 100.0 | 100.0 | 100.0 | 100.5 | 100.0 | 101.2 | 105.3 | **101.2** |
| **1997** | 105.6 | 101.2 | 100.0 | 99.2 | 101.0 | 100.0 | 102.3 | 110.5 | **102.5** |
| **1998** | 108.3 | 102.4 | 96.4 | 98.5 | 101.4 | 100.0 | 102.3 | 110.5 | **102.5** |
| **1999** | 113.9 | 103.0 | 98.1 | 98.2 | 101.9 | 100.0 | 103.5 | 115.8 | **104.3** |
| **2000** | 119.4 | 103.6 | 98.1 | 97.8 | 101.8 | 100.0 | 103.5 | 121.1 | **105.7** |
| **2001** | 125.0 | 104.2 | 100.0 | 97.4 | 101.7 | 100.0 | 103.5 | 126.3 | **107.3** |
| **2002** | 130.6 | 104.8 | 100.0 | 96.9 | 101.6 | 100.0 | 103.5 | 126.3 | **108.0** |
| **2003** | 136.1 | 104.2 | 100.0 | 96.6 | 101.5 | 100.0 | 103.5 | 131.6 | **109.2** |
| **2004** | 141.7 | 103.6 | 100.6 | 96.3 | 101.0 | 106.5 | 104.7 | 131.6 | **110.7** |
| **2005** | 150.0 | 103.6 | 102.6 | 96.3 | 100.4 | 113.8 | 105.8 | 136.8 | **113.7** |
| **2006** | 158.3 | 103.6 | 104.6 | 96.3 | 99.8 | 122.2 | 105.8 | 142.1 | **116.6** |
| **2007** | 162.5 | 103.6 | 108.2 | 96.3 | 99.8 | 122.2 | 105.8 | 142.1 | **117.6** |
| **2008** | 166.7 | 103.6 | 108.2 | 96.3 | 99.8 | 122.2 | 105.8 | 147.4 | **118.7** |

[a] *Key:* 1p = Ratio of childcare spaces to children aged 0 to 5 years of age
2p = Percentage of children doing well on five developmental domains, at age 5 years
3n = Ratio of students to educators in public schools
4p = Average of 5 social and emotional competence scores for 12 to 13 year olds
5p = Basic knowledge and skills index for 13 to 15 year olds
6n = Percentage of PISA scores explained by socio-economic background
7p = Percentage of 20 to 24 year olds in population completing high school
8p = Percentage of 25 to 64 year olds in population with a university degree

[b] Average of education indicators

Figure 7 illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW education domain plus the CIW Domain score for education across the 1994 to 2008 period. The two lines above all others at some distance cross paths in 2001. Up to 2001, the greatest increase occurred in the percentage of the population with a university degree. From 2002 onward, the ratio of childcare spaces to children aged 0 to 5 years showed the highest increases, compared to the other education headline indicators. The bottom line in Figure 7, from 2000 forward, shows that the average of the five social and emotional competence scores for children aged 12 to 13 years had the greatest decrease among all the education headline indicators.

**Figure 7: Trends in Education Indicators, 1994 to 2008**



## *Environment*

Table 8a contains data for eight of 14 headline indicators recommended by Morgan (2011) in the *Environment Domain* for the 1994 to 2008 period. Six are positive and two are negative indicators. They are listed here in the order in which they appear in the headings of the columns of Table 8a: (1n) Ground Level Ozone (population weighted in parts per billion), (2n) Absolute Green House Gas Emissions (megatons of $CO_2$ per year), (3p) Primary Energy Production (petajoules), (4p) Water Yield in Southern Canada ($km^3$), (5p) Viable Non-Renewable Energy Reserves Index, (6p) Viable metal Reserves Index , (7p) Canadian Living Planet Index, and (8p) Marine Trophic Index.

**Table 8a: Trends in Environment Indicators for Canada, 1994 to 2008**

| Year | Headline Indicators for Environment[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1n | 2n | 3p | 4p | 5p | 6p | 7p | 8p |
| **1994** | **36.17** | **624** | **13,913.3** | **1,285** | **1.000** | **1.000** | **1.260** | **3.22** |
| **1995** | **36.03** | **641** | **14,489.2** | **1,270** | **1.613** | **0.987** | **1.250** | **3.15** |
| **1996** | **34.96** | **659** | **14,800.3** | **1,430** | **1.433** | **1.000** | **1.220** | **3.18** |
| **1997** | **36.05** | **672** | **15,284.4** | **1,460** | **1.397** | **0.874** | **1.180** | **3.15** |
| **1998** | **39.44** | **678** | **15,368.7** | **1,200** | **1.443** | **0.814** | **1.110** | **3.13** |
| **1999** | **39.76** | **691** | **15,358.2** | **1,420** | **1.390** | **0.681** | **1.050** | **3.13** |
| **2000** | **34.94** | **717** | **15,768.4** | **1,210** | **1.457** | **0.615** | **1.000** | **3.02** |
| **2001** | **40.66** | **711** | **15,894.9** | **1,200** | **1.507** | **0.593** | **0.970** | **3.06** |
| **2002** | **40.90** | **717** | **16,171.0** | **1,250** | **1.463** | **0.534** | **0.950** | **3.02** |
| **2003** | **39.93** | **741** | **16,170.9** | **1,275** | **1.430** | **0.463** | **0.960** | **3.06** |
| **2004** | **36.22** | **741** | **16,553.7** | **1,335** | **1.480** | **0.508** | 0.960 | **3.04** |
| **2005** | **39.90** | **731** | **16,489.9** | 1,335 | **1.437** | **0.542** | 0.960 | **3.08** |
| **2006** | **37.88** | **718** | **16,815.5** | 1,335 | **1.410** | **0.647** | 0.960 | **3.05** |
| **2007** | **38.73** | **750** | **17,147.9** | 1,335 | **1.607** | **0.627** | 0.960 | 3.05 |
| **2008** | 37.50 | **734** | **16,380.0** | 1,335 | 1.607 | 0.627 | 0.960 | 3.05 |

[a] *Key:* 1n = Ground level ozone (population weighted in parts per billion)
2n = Absolute GHG emissions (megatons of $CO_2$ per year)
3p = Primary energy production (petajoules)
4p = Water yield in Southern Canada ($km^3$)
5p = Viable Non-Renewable Energy Reserves Index
6p = Viable Metal Reserves Index
7p = Canadian Living Planet Index
8p = Marine Trophic Index

* Data which are *not* in bold were obtained by imputation. See text for Table 2a.

After the living standards domain, this domain has the highest percentage of real data, at 88%. On average for the eight domains, 56% of the statistics represented real data and only two of

the eight domains had scores above this average; namely, living standards and the environment. Thus, these two domains are quite a bit ahead of all other domains with respect to their percentage of real data points. Two headliners (i.e., Absolute GHG Emissions and Primary Energy Production) have 100% real data points (all 15) and two more have 14 of 15, or 93.3% (i.e., Ground Level Ozone and Viable Non-Renewable Energy Reserves Index). The Living Planet Index has the fewest real data points (10 of 15, 66.7%), followed by Water Yield in Southern Canada (11 of 15, 73.3%).



**Table 8b: Index of Environment Indicators for Canada, 1994 to 2008**

| Year | Percentage Change in Indicators[a] | | | | | | | | Avg.[b] |
|---|---|---|---|---|---|---|---|---|---|
| | 1n | 2n | 3p | 4p | 5p | 6p | 7p | 8p | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** |
| **1995** | 100.4 | 97.3 | 104.1 | 98.8 | 161.3 | 98.7 | 99.2 | 97.8 | **107.2** |
| **1996** | 103.5 | 94.7 | 106.4 | 111.3 | 143.3 | 100.0 | 96.8 | 98.8 | **106.8** |
| **1997** | 100.3 | 92.9 | 109.9 | 113.6 | 139.7 | 87.4 | 93.7 | 97.8 | **104.4** |
| **1998** | 91.7 | 92.0 | 110.5 | 93.4 | 144.3 | 81.4 | 88.1 | 97.2 | **99.8** |
| **1999** | 91.0 | 90.3 | 110.4 | 110.5 | 139.0 | 68.1 | 83.3 | 97.2 | **98.7** |
| **2000** | 103.5 | 87.0 | 113.3 | 94.2 | 145.7 | 61.5 | 79.4 | 93.8 | **97.3** |
| **2001** | 89.0 | 87.8 | 114.2 | 93.4 | 150.7 | 59.3 | 77.0 | 95.0 | **95.8** |
| **2002** | 88.4 | 87.0 | 116.2 | 97.3 | 146.3 | 53.4 | 75.4 | 93.8 | **94.7** |
| **2003** | 90.6 | 84.2 | 116.2 | 99.2 | 143.0 | 46.3 | 76.2 | 95.0 | **93.8** |
| **2004** | 99.9 | 84.2 | 119.0 | 103.9 | 148.0 | 50.8 | 76.2 | 94.4 | **97.0** |
| **2005** | 90.7 | 85.4 | 118.5 | 103.9 | 143.7 | 54.2 | 76.2 | 95.7 | **96.0** |
| **2006** | 95.5 | 86.9 | 120.9 | 103.9 | 141.0 | 64.7 | 76.2 | 94.7 | **98.0** |
| **2007** | 93.4 | 83.2 | 123.2 | 103.9 | 160.7 | 62.7 | 76.2 | 94.7 | **99.8** |
| **2008** | 96.5 | 85.0 | 117.7 | 103.9 | 160.7 | 62.7 | 76.2 | 94.7 | **99.7** |

[a] *Key:* 1n = Ground level ozone (population weighted in parts per billion)
2n = Absolute GHG emissions (megatons of $CO_2$ per year)
3p = Primary energy production (petajoules)
4p = Water yield in Southern Canada ($km^3$)
5p = Viable Non-Renewable Energy Reserves Index
6p = Viable Metal Reserves Index
7p = Canadian Living Planet Index
8p = Marine Trophic Index

[b] Average of environment indicators

The last figure in the final column of Table 8b reveals a decrease of 0.3% in the average value of the eight environment headline indicators (i.e., in the CIW environment domain index for the period from 1994 to 2008). The last row of the eight distinct indicators in this table reveals that the greatest improvements occurred for the Viable Non-Renewable Energy Reserves Index (60.7%), followed by Primary Energy Production (17.7%). The greatest indicators of

deterioration were the Viable Metal Reserves Index (37.3%), followed by the Canadian Living Planet Index (23.8%). The relatively poor performance on Absolute GHG Emissions (15.0%) is a by-product of the relatively good performance on Primary Energy Production (17.7%), the latter of which is affected by extraction viability as determined by energy prices. This same price-dependency issue affects the Viable Metal Reserves Index. Increased energy production creates more jobs, greater incomes and tax revenues, all of which are good, but in most cases also generates harmful GHG emissions. It is also worth noting that while the Environment Domain Index is comparable to 1994 levels, the historic levels of many of the indicators paint a different picture (see the full *Environment Domain* report for details).



**Figure 8: Trends in Environment Indicators, 1994 to 2008**

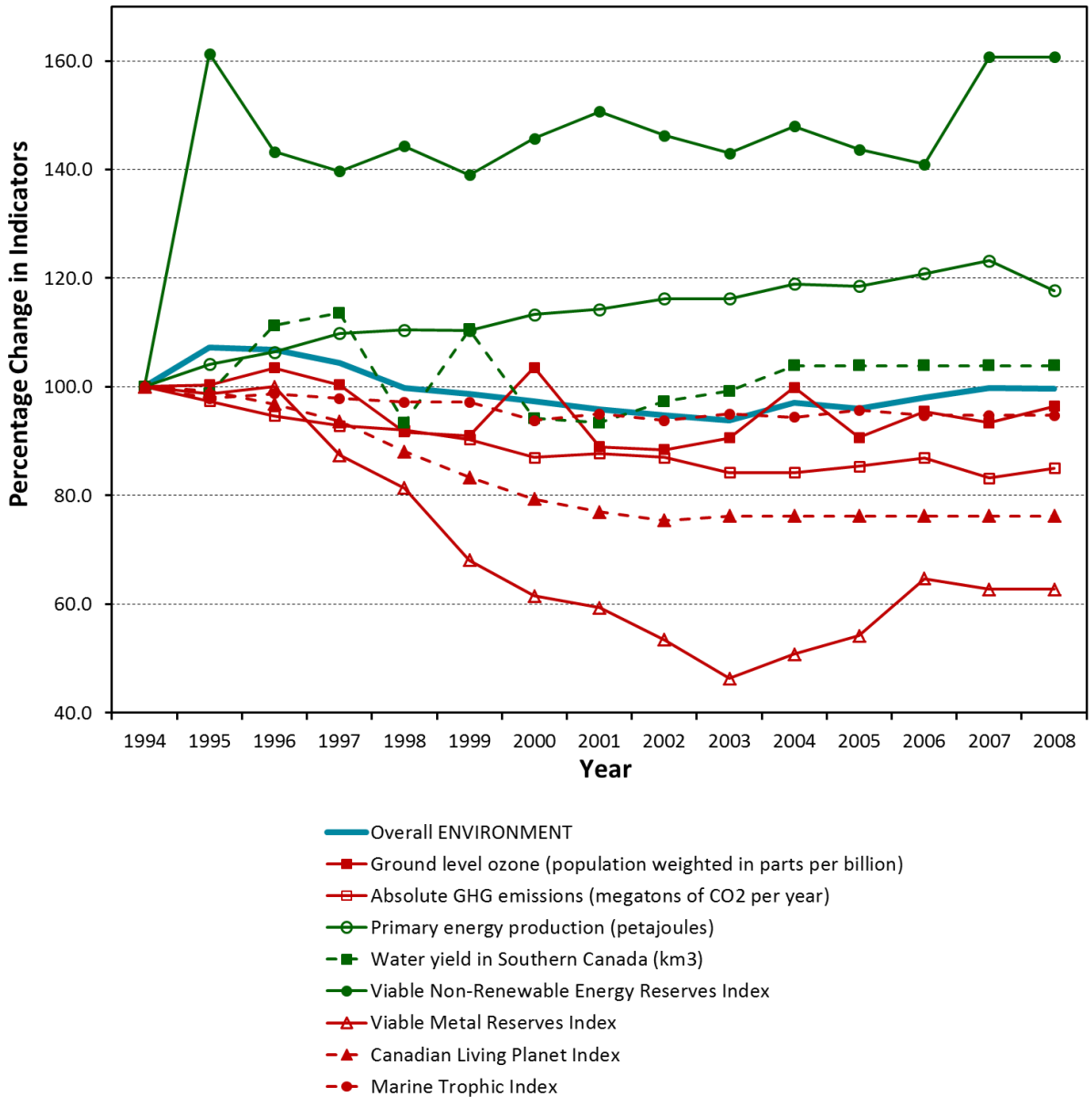


Figure 8 illustrates the trend lines for the average changes for each of the eight headline indicators in the CIW environment domain plus the CIW Domain score for the environment across the 1994 to 2008 period. The line above all others at some distance represents improvements in the Viable Non-Renewable Energy Reserves Index. This is followed by improvements in Primary Energy Production. The line showing the greatest deterioration over the whole period represents the Viable Metal Reserves Index.

Table 9 lists the eight CIW domain index (average) scores. Dividing the eight CIW domain scores by eight, the overall CIW scores were created for the ninth column. The last column of GDP per capita index scores was added for comparison. For the eight-domain CIW, one finds that the index increased by 11.0% over the 1994 to 2008 period. To the extent that such percentage change figures could represent reasonable assessments of the change in Canadians' wellbeing or quality of life over that period, one could say that there was some improvement in our wellbeing. Any comparison of the relatively robust CIW with the very narrow GDP per capita is rough at best. However, the last figure in the last column shows that the picture presented by GDP per capita is rosier than that presented by the CIW (i.e., 31.0% compared to 11.0%). To some extent, of course, the relatively low average percentage change in the CIW is a result of the large number and kinds of imputations required to obtain numerical values for all of the time points.

**Table 9: CIW for Individual Domains and Average of Eight Domains, 1994 to 2008**

| Year | DOMAINS | | | | | | | | Avg. of eight domains | GDP per capita |
|---|---|---|---|---|---|---|---|---|---|---|
| | Living Standards | Healthy Populations | Community Vitality | Democratic Engagement | Leisure and Culture | Time Use | Education | Environment | | |
| **1994** | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **100.0** | 100.0 |
| **1995** | 101.0 | 101.2 | 102.2 | 98.5 | 100.3 | 99.4 | 100.1 | 107.2 | **101.3** | 101.8 |
| **1996** | 100.3 | 102.9 | 103.5 | 96.8 | 100.3 | 98.2 | 101.2 | 106.8 | **101.3** | 102.3 |
| **1997** | 101.6 | 102.9 | 105.6 | 99.2 | 100.2 | 98.0 | 102.5 | 104.4 | **101.8** | 105.6 |
| **1998** | 105.6 | 103.1 | 105.8 | 108.2 | 99.8 | 97.9 | 102.5 | 99.8 | **102.9** | 109.0 |
| **1999** | 110.3 | 100.8 | 105.9 | 117.7 | 99.5 | 97.4 | 104.3 | 98.7 | **104.4** | 114.1 |
| **2000** | 111.4 | 99.2 | 106.2 | 126.9 | 100.3 | 96.9 | 105.7 | 97.3 | **105.5** | 119.0 |
| **2001** | 117.9 | 98.1 | 107.7 | 119.7 | 98.7 | 98.0 | 107.3 | 95.8 | **105.5** | 119.8 |
| **2002** | 115.5 | 100.1 | 109.9 | 115.1 | 97.7 | 98.4 | 108.0 | 94.7 | **105.0** | 122.0 |
| **2003** | 114.4 | 102.7 | 111.1 | 107.8 | 97.8 | 99.2 | 109.2 | 93.8 | **104.6** | 123.2 |
| **2004** | 115.3 | 105.4 | 113.1 | 102.9 | 94.3 | 98.1 | 110.7 | 97.0 | **104.7** | 125.8 |
| **2005** | 116.2 | 108.7 | 114.9 | 111.5 | 96.3 | 98.1 | 113.7 | 96.0 | **107.0** | 128.4 |
| **2006** | 118.5 | 107.1 | 116.2 | 117.9 | 96.9 | 99.0 | 116.6 | 98.0 | **108.9** | 130.7 |
| **2007** | 125.2 | 105.7 | 118.5 | 117.3 | 97.2 | 98.6 | 117.6 | 99.8 | **110.0** | 132.1 |
| **2008** | 126.4 | 106.6 | 120.7 | 119.3 | 97.0 | 99.4 | 118.7 | 99.7 | **111.1** | 131.2 |

Reviewing the last row of Table 9 for the eight domain scores, one finds that living standards improved more than any other domain (26.4%) and, as one would have expected, this result is closer than that for any other domain to GDP per capita. After living standards, the contributors to the overall increase in the CIW are the domains of community vitality (20.7%), democratic

engagement (19.3%), education (18.7%), and healthy populations (6.6%). Only three of the eight domains suffered some deterioration over the whole period and in each case, the level of decrease was low – leisure and culture (3.0%), time use (0.6%), and the environment (0.3%).

**Figure 9: Trends in the Canadian Index of Wellbeing with Eight Domains and Compared with GDP, 1994 to 2008**

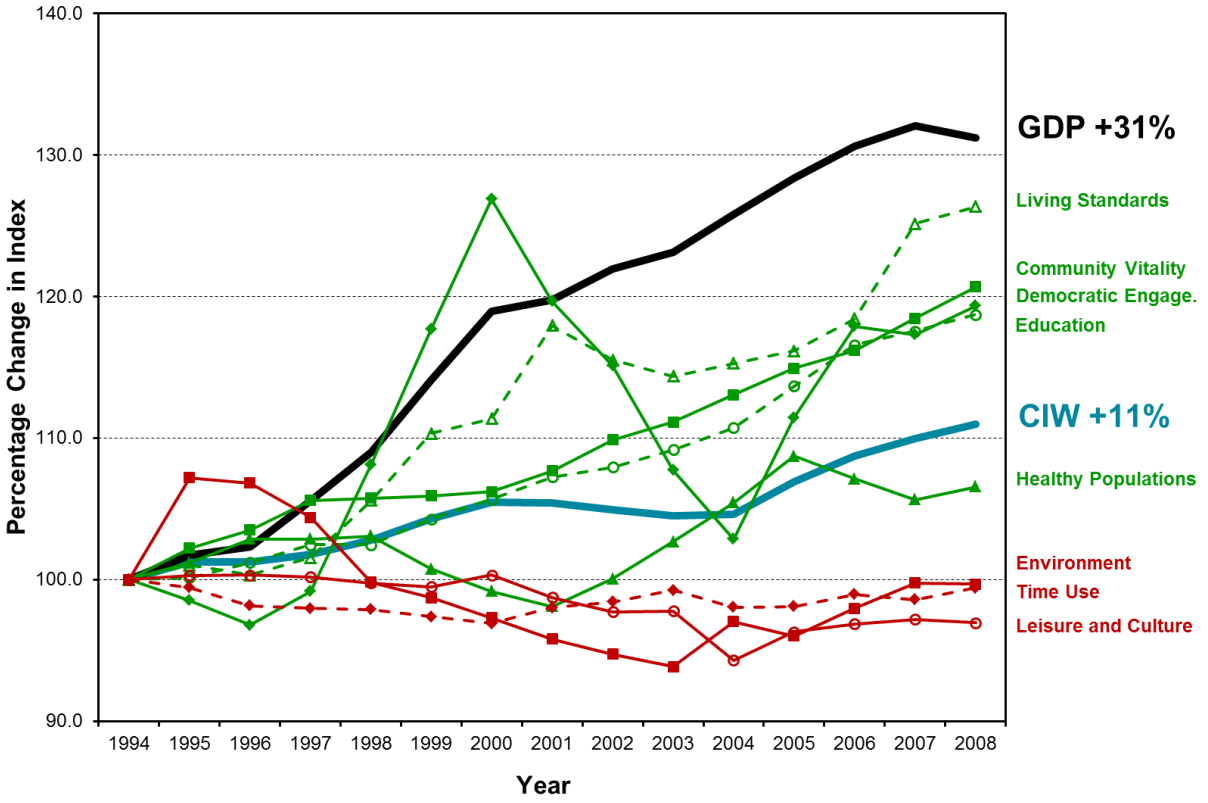



Figure 9 illustrates the trend lines for the 1994 to 2008 period for each of the eight domains, the composite CIW based on those domains, and the Canadian GDP per capita. As indicated in Table 9, the Figure shows that GDP per capita increased more than the composite CIW over the 15 year period. The Figure also shows that the living standards domain trend improved relatively more than all others except that of democratic engagement since 1997. The anomalous spike in the trend of the latter warns us to be cautious in judging its meaning. The four trend lines below the CIW throughout most of the period show that relative deterioration in leisure and culture, time use, the environment, and healthy populations tended to decrease the CIW while improvements in living standards, community vitality, democratic engagement, and education tended to increase it. Assuming all domains are equally important to our wellbeing, we seem to have as much reason to focus on improving the domains where we are relatively strong or on those where we are relatively weak.

Table 10 gives an overview of the final average scores for each of the eight domains, the 64 indicators, and the CIW itself. It provides an opportunity to consider the question of our overall progress over the period from different points of view.

**Table 10**
**CIW List of Indicators for All Domains, with Percentage Gains and Losses, 1994 to 2008**

| *Domain*<br>Indicator | Pct. change[a] |
|---|---|
| ***Community Vitality*** | |
| Percentage reporting participation in organized activities | 27.3 |
| Percentage with 6 or more close friends | 10.1 |
| Property crime rate per 100,000 population | 34.0 |
| Violent crime rate per 100,000 population | 1.1 |
| Percentage who feel safe walking alone after dark | 10.3 |
| Percentage disagreeing that they worry less about the needs of others | 55.6 |
| Percentage who provide unpaid help to others on their own | 15.1 |
| Percentage reporting very or somewhat strong sense of belonging to community | 12.3 |
| **Domain Average** | **20.7** |
| ***Democratic Engagement*** | |
| Percentage of voter turnout at federal elections | -11.8 |
| Percentage that are not interested in politics at all | 36.6 |
| Percentage strongly agree it is every citizen's duty to vote in federal Elections | 14.7 |
| Pct. reporting they are very/fairly satisfied with the way democracy works in Canada | 3.1 |
| Pct. reporting that policies of the federal government have made them better off | 106.6 |
| Ratio of registered to eligible voters | 6.7 |
| Percentage of women in Parliament | 24.4 |
| Net official development aid as a percentage of gross national income | -25.6 |
| **Domain Average** | **19.3** |
| ***Education*** | |
| Ratio of childcare spaces to children aged 0 to 5 years of age | 66.7 |
| Percentage of children doing well on five developmental domains | 3.6 |
| Ratio of students to educators in public schools | 8.2 |
| Average of 5 social and emotional competence scores for 12 to 13 year olds | -3.7 |
| Basic knowledge and skills index for 13 to 15 year olds | -0.2 |
| Percentage of PISA scores explained by socio-economic background | 22.2 |
| Percentage of 20 to 24 year olds in population completing high school | 5.8 |
| Percentage of 25 to 64 year olds in population with a university degree | 47.4 |
| **Domain Average** | **18.7** |
| ***Environment*** | |
| Ground level ozone (population weighted in parts per billion) | -3.5 |
| Absolute GHG emissions (megatons of $CO_2$ per year) | -15.0 |
| Primary energy production (petajoules) | 17.7 |
| Water yield in Southern Canada ($km^3$) | 3.9 |
| Viable Non-Renewable Energy Reserves Index | 60.7 |
| Viable Metal Reserves Index | -37.3 |
| Canadian Living Planet Index | -23.8 |
| Marine Trophic Index | -5.3 |
| **Domain Average** | **-0.3** |

| *Domain*<br>Indicator | Pct. change[a] |
|---|---|
| ***Healthy Populations*** | |
| Percentage self-rated health as excellent or very good | -6.7 |
| Percentage with self-reported diabetes | -49.2 |
| Life expectancy at birth, years | 3.3 |
| Percentage of daily or occasional smokers among teens aged 12 to 19 years | 83.3 |
| Percentage with probable depression | -11.7 |
| Percentage rating patient health services as excellent or good | 2.8 |
| Percentage aged 65 years or more getting influenza immunization | 34.2 |
| Avg. number of remaining years expected to be lived in good health (avg. HALE 15+) | -3.9 |
| **Domain Average** | **6.6** |
| ***Leisure and Culture*** | |
| Average percentage of time spent on the previous day in social leisure activities | -18.5 |
| Average percentage of time spent on the previous day in arts and culture activities | -5.4 |
| Average number of hours in the past year volunteering for culture and recreation organisations | -19.5 |
| Avg. monthly frequency of participation in physical activity lasting over 15 minutes | 24.5 |
| Average attendance per performance in past year at all performing arts performances | 0.5 |
| Average visitation per site in past year to all National Parks and National Historic Sites | -21.4 |
| Average number of nights away per trip in the past year on vacation trips to destinations over 80 km from home | 11.3 |
| Expenditures in past year on all aspects of culture and recreation as a percentage of total household expenditures | 4.2 |
| **Domain Average** | **-3.0** |
| ***Living Standards*** | |
| Ratio of top to bottom quintile of economic families, after tax | -13.9 |
| After tax median income of economic families (2008$) | 24.0 |
| Percentage of persons in low income | 48.9 |
| Scaled value of CSLS economic security | -8.7 |
| Percentage labour force with long-term unemployment | 160.4 |
| Percentage of labour force employed | 8.8 |
| CIBC index of employment quality (1994 QI=100) | -0.8 |
| RBC housing affordability index | -7.7 |
| **Domain Average** | **26.4** |
| ***Time Use*** | |
| Percentage of 20 to 64 year olds working over 50 hours per week | 21.5 |
| Percentage of 20 to 64 year olds reporting high levels of time pressure | -16.3 |
| Percentage of 20 to 64 years old giving unpaid care to seniors | -10.8 |
| Percentage of 65 years and older reporting daily active leisure activities | -4.5 |
| Percentage of 65 years and older reporting annual formal volunteering activities | 9.2 |
| Pct. of 12 to 17 year olds spending two hours or more per day on TV or video games | -14.2 |
| Percentage of 6 to 9 year olds having weekly or more structured activities | 9.0 |
| Percentage of 3 to 5 year olds read to daily by parents | 1.5 |
| **Domain Average** | **-0.6** |

| *Domain*<br>Indicator | Pct. change[a] |
|---|---|
| **Composite Canadian Index of Wellbeing** | **11.0** |

[a] Percentage gains and losses from 100 in base year (1994). Positive values indicate improvement and negative values indicate deterioration.

We learned from Table 9 that there was some improvement in Canadians' wellbeing over the 1994 to 2008 period and that it was the result of improvements in five of the eight domains. From Table 10, one can see that there were 39 headline indicators showing improvements over the period and 25 showing some deterioration. One can also see that four domains had the same number of indicators showing improvement as deterioration (i.e., living standards, healthy populations, leisure and culture, and time use). Two domains had six increases and two decreases (i.e., democratic engagement and education) and only one domain had improvement in every indicator (i.e., community vitality). As well, only one domain had increases in three and decreases in five indicators (i.e., environment).

The total value of percentage increases of the five indicators with the greatest improvements was considerably higher than that of the five with the greatest deterioration (i.e., 477% compared to 137.9%). On the plus side there was the percentage of the labour force with long-term unemployment (159.7%), percentage reporting that policies of the federal government had made them better off (106.6%), percentage of daily or occasional smokers among teens aged 12 to 19 (83.3%), ratio of childcare spaces to children aged 0 to 5 years of age (66.7%), and Viable Non-Renewable Energy Reserves Index (60.7%). On the minus side there was the percentage with self-reported diabetes (49.2%), Viable Metal Reserves Index (37.3%), net Official Development Aid as a percentage of gross national product (25.6%), Canadian Living Planet Index (23.8%), and average visitation per site in past year to all National Parks and National Historic Sites (21.4%).

In all, there were 29 subjective and 35 objective indicators. Among the subjective indicators, there were 20 showing improvement (68.9%), while among the objective indicators there were 19 showing improvement (54.3%).

Practically all participants in the development of the CIW believe that in the interest of the content validity and future development of the index, indicators identified as headliners should not be removed from consideration even though they have severe data limitations. Therefore, every such indicator received full attention in the background studies and some of them are used in the composite CIW. The problem is that trying to keep content validity in terms of the variety of indicators used for each domain, we inevitably lose some content validity in terms of the number of real data points across the 15 year study period; that is, as the variety of indicators increases, the number of imputed statistics required increases. Although this is not the place to investigate the many alternative ways of balancing the trade-offs, it is easy to describe one alternative that favours a reduction in the number of headliners in the interest of a reduction in the number of required imputations.

With potentially 15 real data points for each headline indicator, suppose we remove all indicators from the CIW if fewer than half of their data points have real data (i.e., indicators must have at least 8 of 15, or 53% real data). Table 11 briefly summarizes the results for the CIW. It shows clearly that the somewhat modest requirement for at least 53% real data decreases our headliners from 64 to 31 and increases estimated improvements in wellbeing from 11.0% to 23.7%. Given the depth of the cuts from most domains, one would have serious concerns about the content validity of the remaining set of indicators. What's more, unless there is some change in the definition of most domains, many indicators would be weighted unequally. For example, each headliner in the living standards and environment domains would contribute 1/8 (12.5%) to its domain score, while the single time use score would be 100.0% of its domain score, the two indicators in the community vitality domain would each contribute 50.0% to its domain score, and so on. Unless one believed that it is reasonable to distribute weights unequally in different domains merely on the basis of numbers of available indicators, it is difficult to accept these results.

**Table 11. CIW with Real Data Requirement Set at 53%**

| Domain | Headliners Remaining | Percentage | Pct. Gain/Loss from 100 base |
|---|---|---|---|
| Living standards | 8 | 100.0 | 27.0 |
| Healthy population | 3 | 37.5 | 27.4 |
| Community vitality | 2 | 25.0 | 35.1 |
| Democratic engagement | 2 | 25.0 | -1.2 |
| Leisure and culture | 4 | 50.0 | 18.6 |
| Time use | 1 | 12.5 | 21.5 |
| Education | 3 | 37.5 | 61.4 |
| Environment | 8 | 100.0 | -0.3 |
| **TOTAL** | 31 | | 23.7 |

## Conclusion

We hope that this paper will encourage others to take up the challenge of creating the sort of comprehensive system and composite index envisioned here. Borrowing Churchill's famous remark, we can say that while what we have written here is not the beginning of the end, it is at least the end of the beginning. Although our work on the CIW began some years before the Stiglitz, Sen, and Fitoussi report was proposed, the Commission summarized our own view quite nicely.

> … the commission regards its report as opening a discussion rather than closing it … a global debate around the issues and recommendations raised in this report provides an important venue for a discussion of … what we … care about. … At the national level, round-tables should be established … to identify and prioritise those indicators … the commission hopes that this report will provide the impetus not only for this broader discussion, but for on-going research into the development of better metrics… (Commission, 2009, p. 18)

The absence of adequately resourced specific programs of development of a comprehensive system and composite summary index reduces the chances of researchers raising important questions concerning, for example, the completeness or incompleteness of current stocks of statistical time series, links in the form of causal interactions or mere correlations among the indicators housed in different silos, the collection of redundant indicators needlessly absorbing scarce resources and the failure to collect important data whose availability might reveal serious limitations and/or distortions of our understanding of the quality of our lives, and the consequences of making public policy on the basis of half-baked research (Michalos, 2011).

Given the numbers of active researchers and commentators involved in the project, the number of different tasks that had to be undertaken and the numbers of different ways to proceed, it is understandable that some things appear to be more salient and pressing than others, and that many judgement calls had to be made. Reflecting on our experience, we were reminded of some comments by Fayers and Machin (2007):

> Developing new instruments is a time-consuming task. In summary, our advice is: don't develop your own instrument – unless you have to. Wherever possible, consider using or building upon existing instruments. If you must develop a new instrument, be prepared for much hard work over a period of years. (p. 75)

Of course we agree completely, and we have borrowed and built upon the work of others as much as we could. We now have muddled our way to an illustrative set of domain indexes and an eight-domain composite CIW to help focus attention and provoke discussion and further explorations. There is much more work to be done, but also considerable enthusiasm for getting on with it.

# Appendix 1. Selected Recommendations from Background Studies for Data and Future Work

The recommendations reproduced here have been drawn from the various Domain reports. The report experts were asked to recommend new indicators, domains, procedures, and topics worth considering in the future. This is what they offered.



## Healthy Populations

- Generally, the Healthy Populations domain would be strengthened by improved access to population data stratified by the almost universal predictors of health: age, sex, income, and education. While for some of the core indicators, such as infant mortality, smoking, etc. these covariates were consistently available, for several of the other indicators data were not readily available to enable stratification by income or education. Even when data were available for indicators, they were not always compatible over time (due to varying definitions of categories within indicators, for example) which would allow for clear and accurate interpretations and trends.
- Specific suggestions for future development of this Domain include: (1) a stratified time series analysis of all indicators (both core and secondary) in order to ascertain whether or not the health gap is widening between key population sub-groups over time; (2) inclusion of Aboriginal status as a stratifier; and (3) use of health regions or peer groups to conduct finer regional analysis.
- We are interested in investigating the incidence of specific diseases, but only prevalence data were available. Furthermore, data were not available on potential sub-groupings of interest, such as rural versus urban residence, or disabled versus non-disabled persons.
- We were also limited in the number of data points collected over time. In some cases, only three time points were available, making it difficult to discern true trends over time from irregular "blips" in the data.
- We need data in order to construct a quality-adjusted health expectancy measure such as HALE consistently and repeatedly over time. The disappearance of constituent measures of HALE (health utility index) in recent national surveys present serious limitations to our ability to measure HALE consistently in the future.
- Going forward, we also need a measure that adequately and summarily captures early childhood development across all jurisdictions, and by key stratifiers. We did not include an early childhood development indicator in the current version of this Domain report.
- "Sustainability" is another core value that, like equity, drives the work of the CIW. In future work a core stratified measure indicating sustainability should be incorporated in all analyses.

## Community Vitality

- Further analysis of the initial suite of community vitality indicators is needed, to explore the difference in social relationships and related norms and values for different sub-populations and geographies.
- As in any survey, national level aggregates can and do mask important variation. To this end, it will be critical to improve access to population data stratified by key socio-economic characteristics, including age, gender, ethno-racial identity, household income, educational attainment and disability status.

The community vitality indicators have been selected based largely on the availability of trend data. There are certainly other possible indicators worthy of exploration. Much remains to be known about:

- The types of relationships and networks that exist (e.g., bonding, bridging or linking),
- The purposes of these relationships (e.g., community support, civic engagement and friendship), and
- Their structure (e.g., size, frequency and intensity of contact, density and openness, mobility, power dynamics).

As well, it would be useful to have a better understanding of transactions that occur between people within networks and between organizations (e.g., provision of assistance, exchange of information, application of sanctions).

It will also be important to explore the links between the different domains of the CIW so that we are better equipped to identify and explain notable trends and gaps. More specifically, such an examination in partnership with other organizations working at the local, provincial and national levels, will assist in understanding the impact of positive or negative levels of community vitality – for individuals and communities. This is key to mobilizing people and resources in the pursuit of greater individual and collective wellbeing among Canadians.

## Democratic Engagement

**Thematic Recommendations**

*Individual Engagement*

- Indicators of charitable giving, connecting to religious institutions and connecting with family have been excluded from the Community Vitality Domain and the Democratic Engagement Domain. These indicators have typically been included in social capital models. Consideration may be given to expand one or both of the domains to include these measurements.

*Government Engagement*

- While there are indicators available to measure perceptions of government engagement, there is a lack of consistent and publicly available time-series data on objective measures of government engagement activities. It is recommended that methodologies, indicators, and data sources be developed that measure how governments deliver "governance" and evaluating the results of the process of governing would strengthen this Domain.

*Global Engagement*

- Selecting indicators and obtaining data for measuring government engagement on a global scale is a relatively new area of study that is significantly underdeveloped. The democratic engagement literature review did not produce the clarity anticipated on this item. A specific focus to develop global engagement indicators is recommended, including indicators that are not presently quantified, such as election monitoring.

*General*

- *Thematic relationships:* There is more work required to understand the interrelationship between the themes and indicators that comprise the Democratic Engagement Domain. There is a need to examine potential relationships that exist within and between themes (e.g., would voter participation increase if there were more women elected to parliament and could this in turn affect Canada's commitment to international development aid?).

- *Global relationships:* The literature and existing statistical surveys tend to almost exclusively consider the activities of citizens and governments separately and rarely are they simultaneously considered at a global level. This poses both challenges and opportunities for the continued development of the Democratic Engagement Domain.

- *Impact of engagement activities:* Consideration of the relationship between different models and approaches to democratic engagement and their impact would also be beneficial. For instance, would mandatory voting or a more deliberative democracy change the democratic engagement trends in Canada? Analysis of the successes and challenges of various existing approaches and their application in Canada to improve democratic engagement would be beneficial.

- *Micro considerations:* There is further work required to understand democratic engagement through a disaggregated lens. The analysis in this report considers democratic engagement at the national scale. Collection and analysis of democratic engagement indicators at the provincial and municipal level would be beneficial as would more in-depth analysis of socio-economic characteristics of the indicators by age, gender, income, ethnicity and so forth.

**Broader Recommendations**

- *Internet:* The internet has become a pervasive medium for democratic discussion and information. As important as this information is in adding to our understanding the political voice of citizens and the initiatives of the state to enhance our democracy, there are limited sources of data on internet-based democratic discussion or the extent of our governments' efforts to improve our democracy. It is recommended that research and development of indicators (and methodologies) related to online democratic activities, such as blogging (and other social network sites) and e-government strategies, be pursued. A significant Canadian example is the Canadian Internet Use Survey (CIUS) produced by Statistics Canada.

- *Data Sources:* Should the CIW rely on broad global data sets to measure our national performance or should we instead develop or enhance Canadian sources? The Research Team suggests a combination of both. When focusing on global related indicators, international data sets are helpful in some contexts; however, effort must be made to develop reliable and consistent data sets within Canada at a level of detail useful to the public, policy makers and elected representatives alike.

- *Partnerships and Data Maintenance:* Sustained effort by the Institute and its partners is required to support the continuation of existing sources of democratic engagement data (e.g., the Canada Election Study, the General Social Survey, etc.). Sustained effort to ensure survey questions related to democratic engagement are asked in a consistent manner and that surveys are carried out on a regular basis is necessary to ensure the integrity of the domain over time.

- *Partnerships and Data Maintenance:* There still remains a large gap between the ideal indicators proposed to measure democratic engagement and the availability of data. Of particular benefit would be the administration of a consistent, time-series survey specific to the democratic engagement domain. It is recommended that the domain indicators be included in one survey source in order to maintain consistency, continuity and rigorously examination of changes over time. In the absence of resources to pursue this, establishing a formal partnership with other public or private agencies to conduct and/or analyze surveys containing democratic engagement indicators would be beneficial.

## Leisure and Culture

Even though the recommendations are worded specifically to leisure and culture, the underlying message is that these types of measures would advance the Index in a significant way:

- incorporate more robust measures of *perceptions* related to leisure and culture into existing national surveys of health and wellbeing (e.g., *National Population Health Survey, Canadian Community Health Survey*) would provide critical indicators of the leisure and wellbeing relationship. Simple increases in leisure and culture participation are not always associated with increases in wellbeing; however, increased benefits *perceived* from such participation are often associated with enhanced wellbeing.

- perceptions concerning the extent to which people *value* their free time (i.e., in the *General Social Surveys on Time Use*) should be examined more closely as a potential source for the creation of a composite index beyond the use of single indicators as recommended here.

- data that are available, but not easily accessible, should be located in places that are easy to access, retrieve, and manipulate. For example, because of the emphasis on financial issues in the *Surveys of Service Industries* for both *Heritage Institutions* and *Performing Arts*, the data on attendance to the performing arts and to museums, art galleries, and so on, despite being gathered, are generally not reported in as much detail or in accessible forms or locations. Having access to such data would increase the viability of the measures built on these data and allow for more detailed analyses on subgroups in the population based on, for example, gender and age.

These last two recommendations could be revised to reflect the need for a dedicated, national survey that integrates indicators from *across* the domains to facilitate comparisons and explore interactions/relationships among factors related to our wellbeing.

- a regularly administered national survey dedicated to leisure and culture participation and perceptions would help in alleviating many of the measurement and quality challenges noted earlier. With a focus on leisure and culture, the survey could include: (a) a much broader range of leisure and culture activities (i.e., covering all aspects of arts, culture, and recreation) on which to measure rates of participation; (b) a wider array of composite measures of leisure perceptions based on validated scales (i.e., motivation, satisfaction, perceived benefits, values) as well as opportunities to create new composite measures (e.g., time pressure or time stress indicators built on an expanded set of single-item measures of the perceptions of free time); and (c) a number of measures related to different aspects of health and wellbeing to provide an opportunity to monitor their relationship to leisure over time.

- an exploration of the indicators recommended within other Domains of the CIW project would reveal clear linkages with aspects of leisure and culture and their contribution to the wellbeing of individuals and communities. Identifying these linkages would reinforce the call here for a more comprehensive survey on leisure and culture, either independent of other domains, or integrated into a broader survey, ideally like the *General Social Survey.* For example, leisure researchers typically regard

volunteerism as a form of leisure participation and many forms of social leisure activities as contributing to one's sense of community, social cohesion, and social capital. These perspectives on the nature and contribution of leisure have clear parallels with other Domains in the CIW such as *Community Vitality*, *Healthy Populations*, and *Time Use*.

## Time Use

Specific Recommendations:

- Ensure the continued collection of data pertinent to the working life of Canadians
- Continue to collect data on non-standard work arrangements collected by the GSS
- data on flexible work arrangements need to be enhanced so that they capture whether flexibility is imposed by the employer or chosen by the employee
- Ensure the continued collection of relevant time use data among young Canadians. The Survey of Young Canadians is currently an occasional survey and plans for continued are still tentative. This survey should be continued as it has invaluable information such as screen time and the involvement of extra-curricular activities among children. There is a need for information on children's time spent in family meals with parents or guardians.
- Explore the feasibility of gathering time use diaries for Canadians younger than 15 years (in many European countries time use data are collected for ages 12+
- Enhance existing surveys and time use diaries so that they adequately capture time spent on newer forms of media (such as cell phone texting, etc.)
- Some proposed time use indicators are currently extremely difficult to obtain from Statistics Canada sources. Suggest that these data are made more accessible so that they can be more easily monitored. For instance, time commuting by car is associated with poorer driver health and poorer community health (through increased air pollution). However, these data from the time use diaries are currently very difficult to obtain from the Statistics Canada time use survey so cannot be used as an indicator. As well, there is a need for information of commuting to and from work by mode of transport.
- Time spent caring for an adult or senior with activity limitations is also difficult to obtain and needs to be made more accessible
- Include more robust indicators of time use in existing surveys.

## Education

We make several recommendations for validation research, use of our data (e.g., interpretation), and to interpret CIW data in light of complementary data that allows one to debate social justice and equity issues (e.g., surrounding education for/of/with First Nations

children). The equity and equality issues could potentially be approached better in the future if more data were available to analyze our indicators in relation to other variables, such as immigrant status, English as a second language status, remoteness, socio-cultural background and so on.

## Environment

General recommendations:

- *Improved data collection*: While the Environment Domain had amongst the higher levels of completeness in terms of data, numerous critical gaps exist. Consistent data on water quantity, quality, air quality, biodiversity, land cover, and numerous other areas are absent which restricted the selection of indicators in the Environment Domain Report. Additional support to environmental data gathering agencies is critical to further inform policy decisions.
- *Import-export considerations*: The Environment Domain Report focused primarily upon domestic activity. However, most environmental issues are cross border due to the fact that import and export affect how Canadians affect the planet, and how other countries affect Canada's lands and waters. A more detailed consideration of import-export issues (using input-output tables related to materials, energy, GHGs and water) would strengthen the work.
- *Exploration of linkages*: As the overarching variable (see The Mandala of Wellbeing), it would be interesting to explore the linkages between the Environment Domain and other domains such as Healthy Populations.

Six recommendations for additional indicator areas:

- *Land cover*: Data exist, but little in the way of land-cover change through time. Further efforts could build upon work completed in 2005 by Natural Resources Canada.
- *Material consumption*: Material consumption accounts could be built upon existing economic data (National Accounts), along with the use of the Consumer Price Index to provide consumption baskets. This in turn could ultimately generate a national input-output model of material consumption. This is similar to the notion of the Ecological Footprint, which uses national input-output tables for calculations. The Ecological Footprint could also be reconsidered for use in the future if Statistics Canada felt comfortable with the Global Footprint Network methodology.
- *Toxic chemicals*: At present, toxic chemicals are a notable gap within the indicators. NPRI data are likely the best bet, but there would need to be some kind of weighting applied to the emissions, since toxins have differential impacts that are target-dependent. This work being done in the EU on a chemical risk index which could act as a means of weighting various chemicals based upon toxicity.

- *Food and food security*: Food data are also currently missing from the report. Considerable amounts of data are available on food and agriculture in Canada (e.g., Class 1 prime agricultural lands, organic farming, area under cultivation via Agricultural Census). Statistics Canada also completed a report several years ago on food and the environment, but ultimately food and agriculture were set aside in this report (in part due to the complications of international trade). It could be added in future versions.
- *Water footprint accounting and sustainability*: Water footprinting, which explores import and export of virtual water, as well as sustainable consumption levels within watersheds, would be valuable additions to the CIW. The data largely exist, but were not able to be included in the Environment Domain report due to funding restrictions. Several reports were published after the completion of the Environment Domain report which could inform this in the future and replace the water yield indicator.
- *Ecological representation analysis of protected areas by ecoregion*: Protected areas data are considered important as an indicator (proxy of biodiversity response), but is only really relevant when considered on a habitat representation level by ecoregion. WWF-Canada has done work on this front and such efforts could be conducted on an annual basis in conjunction with CARTS, as it aligns with other biodiversity data reporting efforts such as CESI and the UN CBD.

## Overall System

There is currently a significant push for greater harmonization of official data in Canada. This is evident in the Social Sciences and Humanities Research Council, the Canadian Institutes of Health Research and Statistics Canada. The greater adoption of metadata, especially as envisioned by the Data Documentation Initiative (DDI) should improve this effort. Currently sixty data sets from Statistics Canada that are in the Canadian Research Data Centre Network are being made compatible with DDI through a grant to the Network. These improvements in data documentation could have a very positive impact on the future of the CIW. Still, much more must be done.

If we are going to have comprehensive assessments of human wellbeing or the quality of life, national statistical offices like Statistics Canada must be as aggressive in exploring the full range of subjective determinants and constituents of wellbeing as they are in exploring the full range of objective determinants and constituents. Researchers do not have direct, sensible access to other people's felt affect, attitudes, beliefs, knowledge, motives, values, evaluation standards, needs, wants and personal assessments of happiness or life satisfaction, among other things. So research agendas should include studies of a vast array of phenomena requiring subjective or personal self-reports (i.e., subjective indicators). For examples, self-reported,

- Generic feelings of positive or negative affect
- Specific feelings of fear, joy and contentment
- Attitudes of racism, sexism, classism and homophobia

- Beliefs in progress and democratic process
- Knowledge and appreciation of natural sciences, arts and current events
- Personal standards of evaluation like considerations of equity, justice, moral virtue and beauty
- Motives like vengeance or the pursuit of wealth
- Needs for love, friendship, social and self-esteem
- Expectations and aspirations
- Social comparisons with those worse off or better off than oneself
- Comparisons between one's current status and the best one has ever had or would like to have
- Personal assessments of one's own happiness, satisfaction with life as a whole and with the overall quality of one's own life as well as satisfaction with diverse domains or aspects of life (e.g., job, marriage, housing, family relations) from diverse perspectives
- Personal assessments of one's local community, province or nation as a whole

Explorations of these sorts of phenomena are essential for adequate sustainability assessments because human motives, preferences, needs, perceptions, evaluations and so on can be our greatest resources and/or constraints (Michalos, 1978). The quality of people's lives is a function of the objective conditions in which they live and what they make of those conditions. What they make out of those conditions depends on how the conditions are perceived (accurately or not), what is thought and felt about those conditions (how they are evaluated), what is done and finally, what consequences follow from what is done. Clearly, explorations into these sorts of phenomena will require extraordinary changes in the vision, operation and personnel of national and international statistical offices as they currently exist around the world. Hopefully, the UN declaration of 2005 to 2014 as the Decade of Education for Sustainable Development and the *Istanbul Declaration* were signs that at least some world's offices and agencies are prepared to begin taking on the challenges that such changes imply.

## Appendix 2. Acknowledgements

Over the years of developing the CIW System, a number of people have made diverse contributions and provided assessments, commentary, and advice. Since nobody got everything recommended, it would not be fair to attribute any errors in the text to any of the people listed below. We simply want to express our sincere thanks to all these people for their generous contributions. If any blame is to be laid at anyone's feet, it is the feet of the lead author.

The Atkinson Charitable Foundation provided the initial funding for the research and has continued to be the major contributor to the project since 1999. We can't thank members of the Board enough for their patience, for the trust they put in us, and the opportunity they gave us. Lynne Slotek was the National Project Director for most of the development period, was diligent in guiding and time-managing the work that had to done, and was supremely diplomatic in dealing with the great variety of project participants. The Honourable Roy J. Romanow has been immensely helpful. His political wisdom and stature, his understanding, appreciation of, and commitment to a progressive health care system in Canada as revealed in his official Report of the Commission on the Future of Health Care in Canada (Building On Values, 2002) provided us with a vision and a voice for a better country that we all want to see. We are especially grateful to these members of our team. Below is a list of CIW members, reviewers, and commentators on various aspects of the project.

Hirotsugu Aiga
Mark Anielski
Denis Auger
Pedro Barata
Judith Bartlett
Monique Bégin
Trine Bille
Paul Bowles
Mark Burch
Ye Chen
Andrew Clark
Charles Coffey
Ron Colman
Rupert Crilly
Agnes Croxford
Raymond Currie
Holly Donohoe
Nancy Duxbury
Jane Gaskell
Enrico Giovannini
Steven Goldfinger
Allan Gregg
C. Scott Findlay
Jim Frideres
Michael Frisch
Michael Hagerty
Jon Hall
Ed Hanna
Peter Hardi
Andrew Harvey
Hazel Henderson
Margo Hilbrecht
Nancy Hofmann
Alton Hollett
Thomas Homer-Dixon
Anthony Hutchinson
Magdalena Janus
Sharon Jeannotte
Jochen Jesinghaus
Seymour Kanowitch
Daniel Keating
Noel Keough
Meg Holden
Larry Innes
Marc Lachance
Simon Langlois
Kim Lauzon
Clare Levin
Allen Lomax
Hugh MacKenzie
Filomena Maggino
Massimiliano Mascherini
Doug May
Linda McKessock
Craig McKie
John McMurtry
Robert McMurtry
Terry McRae
Colin Mercer
Hans Messinger
Bill Michelson
John Millar
Byron Miller
Valerie Moller
Alexis Morgan
Dennis O'Farrell
Lars Osberg
Mehmet Ozturk
Charles Pascal
Clem Pelot
Laszlo Pinter
Deborah Poff
Dennis Raphael
Paul Reed
Bill Rees
Rebecca Renwick
Sharon Rich
John Robinson
Roy J. Romanow
Michaela Saisana
Mike Salvaris
David Schindler
Joe Sirgy
Lynne Slotek
Robert Smith
Risa Smith
Leroy Stone
Stefano Tarantola
Dennis Trewin
Charles Ungerleider
Peter Victor
Mathis Wackernagel
Marilyn Waring
Tom Wilson
Susan Wismer
Doug Worts
Sandra Zagon
Jiri Zuzanek
Oliver Zwirner

## References


Andrews, F.M. (1984). Construct validity and error components of survey measures: a structural modeling approach. *Public Opinion Quarterly, 48*, 409-442.

Aristotle. (1999). *Nicomachean ethics*. (2nd ed.). (T. Irwin, Trans.). Indianapolis, IN: Hackett Publishing.

Bollen, K.A. (1984). Multiple indicators: Internal consistency or no necessary relationship. *Quality and Quantity, 18*, 377-385.

Bollen, K.A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*(2), 305-314.

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social Indicators Research, 59*(2), 115-151.

Brooker, A.-S., & Hyman, I. (2010). *Time Use Domain Report*. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Carmines, E.G., & Keller, R.A. (1979). *Reliability and validity assessment.* Beverly Hills, CA: Sage Publications.

CDC. (2000). *Centers for Disease Control and Prevention, Measuring Healthy Days*. Atlanta, GA: CDC.

Commission on Social Determinants of Health. (2007). *Achieving health equity: From root causes to fair outcomes, Interim Statement*. Geneva: World Health Organization.

Commission. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress.* J.E. Stiglitz, A. Sen, & J.-P. Fitoussi (Eds.). Paris: Commission.

Converse, P.E. (1987). Changing conceptions of public opinion in the political process. *Public Opinion Quarterly, 51*, S12-S25.

Cox, D.R., Fitzpatrick, R., Fletcher, A.E., Gore, S.M., Spiegelhalter, D.J., & Jones, D.R. (1992). Quality-of-life assessment: Can we keep it simple? *Journal of the Royal Statistical Society. Series A (Statistics in Society), 155*(3), 353-393.

Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Ed.), *Educational measurement* (2nd ed.) (pp. 443-507). Washington, DC: American Council on Education.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Cummins, R.A. (1996). The domains of life satisfaction: An attempt to order chaos. *Social Indicators Research, 38*(3), 303-328.

Dawes, R.M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34*(7), 571-582.

Dewey, J. (1939). *Theory of valuation*. Chicago, IL: University of Chicago Press.

Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research, 31*(2), 103-157.

Diener, E. (2005). Guidelines for national indicators of subjective well-being and ill-being. *Journal of Happiness Studies, 7*(4), 397-404.

Diener, E., Emmons, R.A., Larson, R.J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71-75.

Diener, E., Lucas, R.E., & Scollon, C.N. (2006). Beyond the hedonic treadmill. *American Psychologist, 61*(4), 305-314.

Diener, E., Sapyta, J.J., & Suh, E. (1998). Subjective well-being is essential to well-being. *Psychological Inquiry, 9*(1), 33-37.

Diener, E., & Seligman, M.E.P. (2004). Beyond money: Toward and economy of well-being", *Psychological Science in the Public Interest,* 5(1), 1-31.

Dunlap, R.E., Van Liere, K.D., Mertig, A.G., & Jones, R.E. (2000). Measuring endorsement of the New Ecological Paradigm: A revised NEP Scale. *Journal of Social Issues, 56*(3), 425-442.

Fayers, P.M., & Hand, D.J. (2002). Causal variables, indicator variables and measurement scales: An example from quality of life. *Journal of the Royal Statistical Society A, 165* (Part 2), 233-261.

Fayers, P.M., & Machin, D. (2007). *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes*. Chichester, UK: John Wiley and Sons.

Gross, B.M. (Ed.). (1967). Goals and indicators of American society. *Annals of the American Academy of Political and Social Science, 371* (Vol. 1, May) and *373* (Vol. 2, September).

Guhn, M., Gadermann, A.M., & Zumbo, B.D. (2010). *Education Domain Report*. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Gurin, G., Veroff, J. & Feld, S. (1960). *Americans view their mental health*. New York: Basic Books.

Hagerty, M.R., Cummins, R.A., Ferriss, A.L., Land, K., Michalos, A.C., Peterson, M., Sharpe, A., Sirgy, J., & Vogel, J. (2001). Quality of life indexes for national policy: Review and agenda for research. *Social Indicators Research, 55*(1), 1-96.

Hagerty, M.R., & Land, K.C. (2007). Constructing summary indices of quality of life: A model for the effect of heterogeneous importance weights. *Sociological Methods and Research*, *35*, 455-496.

Hagerty, M.R., & Land, K.C. (2012). Issues in composite index construction: The measurement of overall quality of life. In K.C. Land, A.C. Michalos, & M.J. Sirgy (Eds.), *Handbook of social indicators and quality-of-life research.* New York*:* Springer.

Horn, R.V. (1993). *Statistical indicators for the economic and social sciences.* Cambridge, UK: Cambridge University Press.

Idler, E.L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior, 38*, 21-37.

IISD. (2009). *BellagioStamp: Sustainability assessment and measurement principles*. Winnipeg, MB: International Institute for Sustainable Development.

Johansson, S. (2002). Conceptualizing and measuring quality of life for national policy. *Social Indicators Research, 58*, 13-32.

Kahn, R.L., & Juster, F.T. (2002). Well-being: Concepts and measures. *Journal of Social Issues*, *58*(4), 627-644.

Kane, M. (2006). Validation. In R.L. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 17-64). Westport, CN: American Council on Education and Praeger.

Kanowitch, S., Michalos, A.C., Slotek, L., & McKessock, L. (2007). *Foundational document*. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Kleijnen, J.P.C. (2001). Ethical issues in modeling: Some reflections. *European Journal of Operational Research, 130*, 223-230.

Kozma, A., & Stones, M.J. (1980). The measurement of happiness: Development of the Memorial University of Newfoundland scale of happiness (MUNSH). *Journal of Gerontology, 35*, 906-912.

Labonté, R., Muhajarine, N., Winquist, B., & Quail, J. (2010). *Healthy Populations Domain Report*. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Land, K.C., Michalos, A.C., & Sirgy, M.J. (Eds.). (2011). *Handbook of social indicators and quality of life research*. New York: Springer.

Lucas, R.E. (2005). Time does not heal all wounds: A longitudinal study of reaction and adaptation to divorce. *Psychological Science, 16*(12), 945-950.

MacCallum, R.C., & Browne, M.W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin, 114*(3), 533-541.

Maggino, F., & Zumbo,B.D. (2012). Measuring the quality of life and the construction of social indicators. In K.C. Land, A.C. Michalos, & M.J. Sirgy (Eds.), *Handbook of social indicators and quality-of-life research.* New York*:* Springer.

Martinez-Alier, J., Munda, G., & O'Neill, J. (1998). Weak comparability of values as a foundation for ecological economics. *Ecological Economics, 26*, 277-286.

McKirahan, R.D. (1994). *Philosophy before Socrates: An introduction with texts and commentary.* Indianapolis, IN: Hackett Publishing.

McMurtry, J. (1998). *Unequal freedoms: The global market as an ethical system.* Toronto, ON: Garamond Press.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp. 33-45). Hillsdale, NJ: Lawrence Erlbaum Associates.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 13-103). New York: American Council on Education and Macmillan.

Michalos, A.C. (1969). *Principles of logic*. Englewood Cliffs, NJ: Prentice-Hall.

Michalos, A.C. (1971). *The Popper-Carnap controversy.* The Hague: Martinus Nijhoff.

Michalos, A.C. (1978). *Foundations of decision-making.* Ottawa, ON: Canadian Library of Philosophy.

Michalos, A.C. (1980). *North American Social Report, Vol. 1: Foundations, population and health*. Dordrecht: D. Reidel.

Michalos, A.C. (1991). *Global report on student wellbeing, Vol. 1: Life satisfaction and happiness.* New York: Springer-Verlag.

Michalos, A.C. (2003). *Essays on the quality of life.* Dordrecht: Kluwer Academic.

Michalos, A.C. (2004). Social indicators research and health-related quality of life research. *Social Indicators Research, 65*(1), 27-72.

Michalos, A.C. (Ed.). (2005). *Citation classics from Social Indicators Research.* Dordrecht: Springer.

Michalos, A.C. (2008a). Education, happiness and wellbeing. *Social Indicators Research, 87*(3), 347-366.

Michalos, A.C. (2008b). Connecting communities with community indicators. In *Statistics, knowledge and policy 2007: Measuring and fostering the progress of societies* (pp. 133-146). Paris: OECD.

Michalos, A.C. (2011). What did Stiglitz, Sen and Fitoussi get right and what did they get wrong. *Social Indicators Research, 102*(1), 117-129.

Michalos, A.C., Creech, H., McDonald, C., & Kahlke, P.M. (2011). Knowledge, attitudes and behaviours. Concerning education for sustainable development: Two exploratory studies. *Social Indicators Research, 100*(3), 391-413.

Michalos, A.C., & Kahlke, P.M. (2010a). Arts and perceived quality of life in British Columbia. *Social Indicators Research, 96*(1), 1-39.

Michalos, A.C., & Kahlke, P.M. (2010b). Stability and sensitivity in perceived quality of life measures: some panel results. *Social Indicators Research, 98*(3), 403-434.

Michalos, A.C.,& Robinson, S.R. (2011). The good life: Eighth century to third century BCE. In K.C. Land, A.C. Michalos, & M.J. Sirgy (Eds.), *Handbook of social indicators and quality-of-life research.* New York*:* Springer.

Michalos, A.C., Ramsey, D., Eberts, D., & Kahlke, P.M. (in press). Good health is not the same as a good life: Survey results from Brandon, Manitoba. *Social Indicators Research*.

Moore, K., Swystun, L., Holden, B., Bernardin, H., Dunning, B., & Graham, P. (2010). *Democratic Engagement Domain Report*. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Morgan, A. (2011). *Environment Domain Report*. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Munda, G. (2004). Social multi-criteria evaluation: Methodological foundations and operational consequences. *Journal of Operational Research, 158*, 662-677.

Munda, G. (2006). Social multi-criteria evaluation for urban sustainability policies. *Land Use Policy, 23*, 86-94.

Munda, G., & Nardo, M. (2005). *Non-compensatory composite indicators for ranking countries: A defensible setting*. Ispra, Italy: European Commission, Institute for the Protection and Security of the Citizen.

Nardo, M., *et al*. (2005). Handbook on constructing composite indicators: Methodology and user guide. *OECD Statistics Working Papers*, 2005/3, OECD Publishing.

Nunnally, J.C. (1967). *Psychometric theory.* New York: McGraw-Hill.

OECD. (2007). *Istanbul Declaration.* Paris: OECD.

Okun, M.A., & Stock, W.A. (1987). The construct validity of subjective well-being measures: An assessment via quantitative research syntheses. *Journal of Community Psychology, 15*, 481-492.

Osberg, L., & Sharpe, A. (1998). *An index of economic wellbeing for Canada*. Research Paper R-99-3E. Applied Research Branch, Human Resources Development Canada, Ottawa, ON.

Oud, J.H.L., & Delsing, M.J.M.H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Monfort, J.H.L Oud, & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp.201-244). Heidelberg: Springer.

Ravallion, M. (2010). *Mashup indexes of development*. Washington, DC: The World Bank.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research, 81*(1), 65-77.

Scott, K. (2010). *Community Vitality Domain Report.* Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Sen, A. (1985). Well-being, agency and freedom: The Dewey lectures 1984. *Journal of Philosophy, 82*(4), 169-221.

Sen, A. (1999). *Development as freedom.* New York: Random House.

Sharpe, A., & Arsenault, J.-F. (2009). *Living Standards Domain Report.* Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Shepard, L.A. (1993). Evaluating test validity. *Review of Research in Education, 19*, 405-450.

Sirgy, M.J., Michalos, A.C., Ferriss, A.L., Easterlin, R.A., Patrick, D., & Pavot, W. (2006). The quality-of-life (QOL) research movement: Past, present and future. *Social Indicators Research, 76*(3), 343-466.

Smale, B., Donohoe, H., Pelot, C., Croxford, A., & Auger, D. (2010). *Leisure and Culture Domain Report.* Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

UNESCO Education Sector. (2006a). *Framework for the UNDESD International Implementation Scheme.* Paris: UNESCO.

UNESCO Education Sector. (2006b). *Education for Sustainable Development Toolkit.* Paris: UNESCO.

Victor, P.A., Hanna, J.E., & Kubursi, A. (1995). How strong is weak sustainability? *Economie Appliquée, 48*(2), 75-94.

Wolming, S., & Wikstrom, C. (2010). The concept of validity in theory and practice. *Assessment in Education: Principles, Policy and Practice, 17*(2), 117-132.

World Commission on Environment and Development. (1987). *Our common future.* New York: Oxford University Press.

Wright, J.G., & Feinstein, A.R. (1992). A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *Journal of Clinical Epidemiology*, *45*(11), 1201-1218.

Zeller, R.A., & Carmines, E.G. (1980). *Measurement in the social sciences: The link between theory and data.* Cambridge: Cambridge University Press.

Zumbo, B.D. (1998a). Opening remarks to the special issue on validity theory and the methods used in validation: Perspectives from the social and behavioral sciences. *Social Indicators Research, 45*, 1-3.

Zumbo, B.D. (Ed.). (1998b). Validity theory and the methods used in validation: Perspectives from the social and behavioral sciences. Special Issue, *Social Indicators Research, 45*(1), 1-509.

Zumbo, B.C. (2007). Validity: Foundational issues and statistical methodology. In C.R. Rao & S. Sinharay (Eds.)*, Handbook of statistics* (Vol. 26) (pp. 45-79). The Netherlands: Elsevier Science.

Zumbo, B.D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R.W. Lissitz (Ed.), *The concept of validity: Revisions, new directions and applications* (pp. 65-82). Charlotte, NC: Information Age Publishing.

Zumbo, B.D., & Forer, B. (2011). Testing and measurement from a multilevel view: Psychometrics and validation. In J.A. Bovaird, K.F. Geisinger, & C.W. Buckendahl (Eds.), *High stakes testing in education – Science and practice in K-12 settings* (pp. 177-190). Washington, DC: American Psychological Association Press.

## *For More Information*

To find out more about the Canadian Index of Wellbeing, please go to www.ciw.ca or e-mail info@ciw.ca.

Measuring what matters

Based in the Faculty of Applied Health Sciences at the University of Waterloo, the Canadian Index of Wellbeing Network is an independent, non-partisan group of national and international leaders, researchers, organizations, and grassroots Canadians. Its mission is to report on quality of life at the national level and promote a dialogue on how to improve it through evidence-based policies that are responsive to the needs and values of Canadians.

The Network's signature product is the Canadian Index of Wellbeing (CIW). The CIW measures Canada's quality of life and tracks progress in eight interconnected categories. It allows us, as Canadians, to see if we are better off or worse off than we used to be — and why. It helps identify what we need to change to achieve a better outcome and to leave the world a better place for the generations that follow.

The Honourable Roy J. Romanow, Chair
The Honourable Monique Bégin, Deputy Chair

WATERLOO APPLIED HEALTH SCIENCES | CANADIAN Index OF WELLBEING

University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1
519-888-4567, ext. 31235 | ciwinfo@uwaterloo.ca | ciw.ca