

Summary of the Conference
“Review of the Child Well-being Index”

co-sponsored by

Brookings Center on Children and Families and Foundation for Child Development

May 10, 2006

Washington, D.C.

Summary written by Ron Haskins and Howard Rolston

This document summarizes four background papers and subsequent discussion of each paper at a conference on the Child and Youth Well-being Index (CWI). The conference, co-sponsored by the Brookings Center on Children and Families and the Foundation for Child Development (FCD), which has also sponsored the development and annual publication of the CWI by Ken Land and his team at Duke University, was organized to provide critiques of several conceptual and technical issues raised by the CWI. The four topics addressed at the conference, and the authors of the respective background papers, were as follows: how the CWI measures disparities between ethnic groups (Donald Hernandez), state and local (as opposed to federal) measures of well-being (William O’Hare), whether the CWI should use weights to express the relative importance of various domains of child well-being (Nick Zill), and whether the specific domains of well-being used by the CWI are representative of available information about child and youth well-being (Brett Brown). In addition, conferees had an interesting discussion of ways to bring public attention to child well-being in general and the CWI in particular. This discussion was initiated by brief presentations by Harold Leibovitz of FCD and Melissa Skolfield of Brookings. The conference was attended by thirty-six scholars, advocates, foundation executives, students, and staff of federal policymakers.

Measuring Social Disparities: Ethnic, Racial, SES, and Immigrant Status; Donald Hernandez, State University of New York, Albany

Summary of Paper

The CWI computes ethnic differences relative to a base year within each ethnic group. For each of its 28 indicators scores and seven domain scores, annual scores are computed relative to the initial or base year. These base scores are computed separately within ethnic groups for each indicator and then indicators are averaged within domains to produce the scores in each of the seven domains. The domain scores are then averaged to produce the overall score on the CWI. Each ethnic group begins with a score of 100; changes in subsequent years are expressed as percentage changes from that group’s score in the base year. If whites scored 50 on a given indicator or domain in the base year, their CWI score would be assigned the value of 100. Similarly, if Hispanics scored 40 in the base year, they also would be assigned the score of 100 for that year. Thus, although whites and Hispanics had a 10 percentage point difference in scores, this absolute difference in performance is masked by the way scores are computed. The same procedure is followed each year, with scores for each group being computed relative to its own score in the base year. This method of computing scores produces a reasonable

picture of changes over time within and between ethnic groups, but it does not measure absolute differences between groups.

Don Hernandez developed a scoring procedure that captures the absolute differences in scores between whites, blacks, and Hispanics. Specifically, scores in the base year are computed across all three ethnic groups and a single average is computed. This average score is then assigned a value of 100 and each ethnic group's score is computed for the base year relative to the group average. If whites scored 10 percent higher than the overall average in the base year, their score would be computed at 110. If Hispanics scored 10 percent lower, their score would be computed at 90.

Next, Hernandez calculates the domain and overall CWI score for subsequent years separately for each ethnic group by averaging that group's scores over the indicators within each domain and across the seven domains for the overall CWI score. However, Hernandez computes the scores within each group relative to the population as a whole in the base year. If in a given year whites scored 5 percent above the population average in the baseline year, their score for the year would be 105; if Hispanics scored 5 percent below the overall average, their score would be 95. Based on this scoring procedure, we could conclude that Hispanics had a gap of 20 points in the baseline year, but subsequently narrowed the gap to 10 points. Thus, the Hernandez scoring method shows both trends in scores as well as absolute differences between the ethnic groups in all years.

Using his scoring method, Hernandez proceeds to analyze changes in domain scores and the overall index, with an emphasis on Hispanic-white and black-white disparities. In the baseline year of 1985, whites score higher than both Hispanics and blacks on every domain except the emotional/spiritual domain in which both blacks and Hispanics scored higher than whites. Hernandez also finds that Hispanics closed the gap with whites on the overall index score and on six of the seven domains (all except community connectedness). Similarly, blacks closed the gap with whites on the overall index score and on five of the seven domains (all except social relationships and emotional/spiritual).

Comments by Participants

1. Many participants observed that the CWI method of scoring disguises differences between ethnic groups. The CWI captures the contrast in patterns of change over time between ethnic groups, but because scores for each ethnic group are expressed relative to the base year within each group, the absolute differences between groups are lost. There was widespread agreement that the Hernandez approach that made the absolute differences between ethnic groups transparent was a valuable supplement to the CWI.
2. The education domain in the CWI is limited because some important indicators of educational progress are not used to determine the score. The dramatic strides in college graduation rates since 1975 by black females, for example, would not be reflected in the score for blacks on the education domain because college graduation rates are not included in the CWI. Some participants believed that the CWI should include high school graduation rates and college attendance rates.
3. Some measures of educational progress are flawed. There is lots of evidence, for example, that high school graduation rates are inflated by local school systems. If

- the measures of education in the CWI were expanded, some believed that it would be unwise to include unreliable measures such as high school graduation rates.
4. In a comment that came up during several of the sessions, participants expressed the concern that reducing child well-being to a single score masks lots of information such as differences between ethnic groups in particular domains. If blacks made progress in college attendance while whites remained flat but the income gap between blacks and whites increased, these contrasting patterns could cancel each other out in the computation of the overall index score, losing all information about the relative differences in changes between blacks and whites. One suggestion to deal with this issue was to present lots of information about changing patterns on indicators and domains for each ethnic group. Ken Land pointed out that their website already presents these types of comparisons.
 5. A valuable supplement to the CWI would be analyzing trends by the socioeconomic status (SES) of children. The problem is that many of the CWI indicators do not have SES measures such as education or income of parents. Don Hernandez plans to use SES measures in the analysis of a few of the measures for which SES information is available.
 6. One participant commented that the survey that yields most of the indicators for the CWI's safety domain (smoking, drinking, drug use) is given only to students in school and ignores dropouts and incarcerated individuals. Presumably, the result would be to understate the estimates of risky behaviors.
 7. One observer commented that a number of the self-report measures in the CWI are flawed because of underreporting.
 8. The safety domain in the CWI would be strengthened if childhood accidents were included.
 9. It would be useful to perform statistical tests to determine whether some of the CWI trends represent significant changes over time or whether differences in trends between ethnic groups are significant. It is suspect to simply report the level of changes without testing to determine whether the changes are merely random.

State-Level Indicators, William O'Hare, Annie E. Casey Foundation.

Summary of Paper

The current CWI reflects child well-being at the national level. Based on his long experience directing the Annie E. Casey Foundation's KidsCount, Bill O'Hare's paper describes the potential and the limitations of extending a CWI-type index to the state-level or below. He emphasizes the increasing importance of states as the locus of policymaking that affects children and, based on data from KidsCount, the enormous variation in child well-being across the states. On ten KidsCount measures the number of states with a significantly different estimate from the national estimate varies from 19 to 49. Of 500 KidsCount state measures, 340 are statistically different from the national average.

Data at the national level are currently not available for many measures at the state level and certainly not back to 1975, the earliest year for which the CWI can be calculated. Even less data are available at the sub-state level. However, there is growing

potential for state-level, and in some case county or city, statistics on child well-being. Some of the key new or expanded data sets include: the Census Bureau's American Community Survey (ACS), the National Survey of Children's Health derived from the State and Local Area Integrated Telephone Survey (SLAITS), and the National Assessment of Educational Progress (NAEP), now mandated for all states by the No Child Left Behind law.

Key limitations of using data at the state level are small sample sizes and lack of consistency in definitions both across states and over time within states. The small numbers are both a problem in the number of individuals sampled, but also in some cases in the number of actual events for some measures. For example, in many states there are so few teen suicide deaths that yearly percent changes that are huge (300 percent or more) simply reflect random variation. A vivid example of state variation in definition is the enormous differences in what many states report regarding the educational proficiency of their students in comparison to what NAEP reports.

In communicating results, there are many different ways that data for states could be arrayed, including highest to lowest individually or in clusters, regionally, or on a map. Having too many measures can be difficult to communicate to the media and public, but the use of domains such as the seven domains used in the CWI may not be feasible at the state level, given that fewer measures are available and that some domains may be empty. Ultimately, some of these choices would depend on the primary audience that FCD would want to reach with a state-level CWI.

Comments by Participants

1. There was discussion of the trade-offs between limiting presentation of state differences to description and including analysis of the differences. Some argued that going beyond description would create controversy and deflect from the main purpose of describing where states stand, whereas others felt that it was preferable to have informed analysis rather than receivers of the information doing their own analysis, whether informed or not. Those who favored simple description of state-level data suggested that, if analysis is considered necessary, description and analysis occur in different venues or that demographic data be presented in addition to the child well-being measures as KidsCount now does on opposite pages. This approach could also help clarify where compositional changes are contributing to changes in outcomes. One of the participants calling for more analysis cited a study that found that race and income account for 75 percent of the variation in state KidsCount measures, suggesting that including no analysis could be misleading. Another participant argued for a social indicators model that explicitly links policy to outcomes.
2. Differing views were expressed on the utility of including domains and the trade-off between a few "powerful" measures and a larger number of measures that were more comprehensive. Some argued that fewer measures were less confusing to communicate and were particularly relevant at the state level where fewer measures are available. Others argued that the use of domains helps to bring the data to the attention of disparate policy communities. There were suggestions that rather than focus on domains, other ways of unifying measures might be attempted, such as integrating those that contribute to an important positive adult

- outcome. In this vein, one participant proposed trying to develop a model comparable to the public health concept of “disability-free life expectancy.” [Editors note: This point applies also to the Zill paper on weighting.]
3. Some participants stated that often it was necessary to get below the state level to capture the attention of policymakers, and that there are often within-state differences that are as large as across-state differences.
 4. One participant stated that an important feature of the American Community Survey (a relatively new Census Bureau survey that is the biggest survey of its type ever conducted) is that it would allow the identification of outcomes for subgroups of children, and particularly low-income children, who might be of most concern and most affected by social policy.
 5. Another participant suggested that the current way that the CWI is computed using year-to-year percent changes biases the index upward. For example, an increase of 5 to 10 on a measure is a 100% increase, whereas the inverse reduction from 10 to 5 is only a 50% decrease. This might be more of a problem at the state level. [Editor’s note: This point is also relevant to Zill’s paper on weighting.]
 6. It was noted that the only data that are comparable across states are those that come out of the federal statistical system. There is limited incentive for states to increase cross-state comparability of administrative data.
 7. It’s rarely the case that year-to-year changes within states are statistically significant, so there would be little reason to report them. There is little interest among the media in significance either over time or across states. Bill O’Hare said that despite the fact that KidsCount has significant cross-state information available, it has only rarely been requested.

Using Weights to Express the Relative Importance of Specific Domains in the Overall Index Score, Nick Zill, Westat.

Summary of Paper

In computing the current CWI, the measures in each domain are equally weighted to produce a score for the domain, and then the seven domains are equally weighted to produce the overall index score. Land and his colleagues have argued that equal weighting is simple, transparent, and replicable, and that this approach achieves the greatest level of consensus in the absence of a clear ordering of the individual measures.

Nick Zill’s paper raises questions about equal weighting and suggests alternative weighting approaches. Zill argues that equal weighting raises both causal and perceptual issues. The former is problematic insofar as some components may have greater significance than others for child well-being, and the latter insofar as public perception likely would not afford all measures equal weight. Zill cites a large body of data indicating that the health of children has improved substantially over the last quarter century, including sharp declines in infant and child mortality across the entire age range from infancy to adolescence. Yet the CWI health domain shows a sharp decline in children’s health due to the large increase in child obesity.

Zill describes three possible alternative weighting strategies: factor analysis, scaling, and regression analysis of longitudinal data. Factor analysis seeks to explain the

variability of a set of observed variables such as those that comprise the CWI in terms of a smaller number of unobservable “factors.” Zill and colleagues have performed a number of analyses using KidsCount data and the results showed that the variability of a large number of measures could be explained by a single factor. This finding is supportive of the equal-weighting strategy of the CWI. However, Zill also identified two additional independent factors implying the inadequacy of equal weighting, and Zill suggests that the existence of these additional factors is worthy of further exploration.

Scaling, as it would apply to the CWI weighting of measures, could involve collecting information from child development experts or the public on the relative importance of the components of the CWI for child well-being. Statistical techniques would then be applied to these judgments to create a set of scale values. If the values were all quite similar, the CWI’s current equal weighting approach would be supported. But if the values expressed regarding the relative importance of CWI components were different, the scaled values could be used to create weights for the composite CWI. One shortcoming of this approach is that since these data are not currently available, it would be necessary to conduct surveys to gather the information needed to create the scales.

Regression analysis would analyze longitudinal data relating how children scored on the measures included in the CWI to positive adult outcomes using a data set such as the 1979 National Longitudinal Survey of Youth or the National Education Longitudinal Study (NELS). An adult measure could be whether the child has grown into a gainfully employed, self-sufficient, and healthy adult at age 30. The coefficients from the regression would be used to create weights for the CWI, reflecting their causal contribution to the adult outcome(s). This approach would make an explicit assumption that the CWI is measuring child well-being for the purpose of determining how well children are being prepared to take on adult responsibilities rather than their subjective sense of well-being as a child. The approach also assumes that such a regression is not misspecified and therefore accurately reflects the causal influence of each measure.

Comments by Participants

1. Although the CWI is equally-weighted (measures within domains and domains within the composite index), a number of participants commented that, given the variety of different types of measures, along with more or less arbitrary choices regarding how they are formulated, the measures are implicitly non-equally weighted. Some values are rates and others are mean scores so that percent change from the base year (the way CWI scores are now calculated; see above) can mean different things. For example, since childhood obesity was 5 percent in the base year and 15 percent more recently, it has increased by 200 percent, but a National Assessment of Educational Progress (NAEP) score cannot increase or decrease by 200 percent even if very large changes occur. The upshot is that obesity has a much bigger implicit weight in the current index than NAEP scores. In the same vein, another commentator made the point that if the mode of presenting information on the weight status of children was formulated as a “normal weight” measure, it would have decreased from 95 percent to 85 percent, a change of only about 9 percent, rather than the current 200 percent, giving it a much smaller weight, despite the fact that the same information was included in the measure. Finally, the fact that some domains have more measures than others

- implicitly weights measures in the latter more highly. Some argued that the matter should be examined empirically and that Nick Zill's analysis supported equal weighting. Others said that the current weighting method which is equally weighted formally was best because it was simplest and most transparent.
2. In the context of the above discussion, it was noted that the health domain would have shown improvement, instead of substantial worsening if the obesity measure were removed. This observation led to an extended discussion of how much the increase in obesity among children contributes to a decline in their health status. Some argued that life expectancy among adults has increased despite a very substantial increase in adult obesity and obesity and that childhood obesity was not strongly linked to adult obesity. Others argued that the former was highly predictive of the latter and that lifetime obesity had more negative effects on health than obesity acquired later in life.
 3. Ken Land responded that he and his colleagues had done a great deal of methodological work which supported the current method of equal weighting. This work included establishing that trends in high school seniors' responses to subjective well-being questions on the Measuring the Future Survey tracked the CWI closely. In addition, he cited his and others' work showing that in the absence of strong agreement on weights of different measures, equal weighting would produce the highest levels of consensus. Finally, he said that weighting each measure individually made very little difference compared to the current method of weighting all domains equally. He did not directly address the comment that the CWI was implicitly not equally weighted within domains.
 4. Several participants supported the third strategy in Nick Zill's paper which would seek to establish empirical links between the CWI measures and a dominant normative goal as identified by interviews with experts or the public. Another possibility for a dominant normative goal would be an intrinsically meaningful index, such as "projected life expectancy." One commenter said that the current CWI implicitly assumes norms, but Ken Land argued that this was not necessarily so, and that the CWI was rooted in empirical research.
 5. There was extensive discussion of having fewer or more measures. One commentator suggested greatly increasing the number of measures because that would reduce the importance of weights. Another said that fewer measures would be easier to communicate than more, but would eliminate domains. Some questioned the value of domains, but another participant reiterated their policy meaning, for example, their linkage to Congressional committees.
 6. One participant questioned Nick Zill's factor analysis, noting that there were many measures and few states. In response to the assertion that there were many data points taking into account the multiple observations over time, the participant argued that these observations were not independent.

Does the CWI Measure Representative Domains of Child Well-being? Brett Brown, Child Trends.

Summary of Paper

Brett Brown's paper does not explore the question of whether the CWI measures representative domains by examining and critiquing the particular seven domains of the CWI. Rather, Brown stated that the CWI does about as well as can be done given that it's tethered to data that have been collected since the 1970's. Instead of a detailed focus, Brown takes a macro-level, strategic view and recommends changes to the CWI based on emerging, richer data and research on child development and well-being. His approach is intended not only to reflect recent improved data collection, but also to suggest how the CWI might respond to the certainty that data quality and quantity will continue to improve in the future.

Although the 30-year CWI has been useful in establishing an historical perspective, a rolling 10-year index would be more than adequate for policy considerations and would allow continuous improvement of the CWI as richer measures become available. This ten-year index would be supplemented with a "state-of-the-art" index which would be based on the best available data at the time. It would be useful for comparing subgroups of children, but not for trend analysis. Finally, an "ideal" index would be developed based on the strongest constructs that research on children has produced, even if data were not yet available to calculate it. This approach would help to guide future data and index development.

The current CWI includes both direct measures of child well-being and measures of children's environment which research suggests influence well-being. It would be useful to separate these into two indices, especially since policy can primarily affect environments.

In addition, the current CWI is based on a Quality of Life (QOL) framework which is adequate to summarizing the limited set of indicators that has been available from the 1970's. A virtue of the QOL is that it's a general model applicable across all ages. Since the 1970's, however, a great deal of research has been rooted in the developmental/ecological framework which identifies periods of development that are connected to milestones and challenges particular to each period. The research conducted under this model has been the basis for many of the rich, new measures of child well-being that have been developed. The CWI should explicitly adopt the developmental model as its guiding framework.

Comments by Participants

1. A number of participants supported the idea of a ten-year CWI that would include newer measures. Although Brown put this forward as a substitute for current CWI, many argued for it as a supplement. One commenter suggested that using the ten-year CWI as a supplement would be parallel to the Census Bureau's development and publication of alternative poverty measures. One participant said that a virtue of the longer time span is that reporters were frequently interested in how a given year compared to the highest or lowest year.
2. Ken Land said that his team was looking into separating outcome measures from environmental measures. One participant presented examples of how it was difficult to define "outcome" unambiguously, and that what counted as an outcome in one context, didn't in another. Another person commented that direct measures of children qualify as outcomes.

3. Brett Brown argued that if FCD didn't begin to incorporate newer, richer measures, he believes that other organizations will begin to create other indices which do. Ken Land said that many of the new measures that Brown mentioned were not available annually. Others argued for the importance of incorporating new information even if it required some accommodation.
4. There was discussion about the fact that the current CWI includes little on younger children – most of the measures focus on teens. Vicki Lamb acknowledged this focus but said that they were looking at age breaks and had included some additional measures appropriate to younger children in an extended CWI. Ken Land mentioned that the Monitoring the Future survey was now including a sample of 8th and 10th graders, and that a 5th grade sample could be added.
5. Ken Land and Vicki Lamb both said that current measures of emotional well-being are weak and that the current social relationship measures are very abstract. One participant suggested that teachers' ratings of children's behavior are well validated and would be a good way to add socio-emotional measures. The same participant also suggested adding background measures to the NAEP, but Brett Brown said that the authors of the NAEP have already removed many of them as "superfluous."
6. Another participant said that the National Institute of Child Health and Human Development was for the first time directly funding development of measures and although NICHDDP funding is not directed at population measures might support population measures in the future.
7. Another commenter raised the question of what is known about the number of children who are incarcerated and how incarceration would affect the CWI, since most of the measures are based on household surveys which exclude incarcerated individuals.

Bringing Attention to the CWI and the Status of Children: A Lunch Discussion

Summary of Harold Leibovitz's Remarks

The Foundation and its Board supported the development of the CWI to create a single measure of overall child well-being, and to bring as much attention as possible to changes in the status of children over time. FCD believes that the simplicity of a single measure of child well-being – like the poverty index and the unemployment rate – can provide a straightforward and easily understandable way of expressing child well-being and trends in child well-being. In addition to the single score, the CWI's domain scores highlight specific areas of child well-being that can help policymakers identify areas of policy success as well as areas that need attention.

The Index has been released each year for the past three years in a public event at Brookings. Each year the Foundation has selected one of the domains for more in-depth analysis. The domains examined in the annual release events include obesity, crime, and education. Each year FCD has hired a Washington-based communications firm to help publicize the release, the result of which has been extensive media coverage of both the index score and the domain selected for detailed examination.

Leibovitz concluded by observing that the Foundation is pleased with its progress in publicizing the CWI, but Foundation officers want to consider innovative ways to bring more attention – especially of the public and of policymakers – to the CWI and to children’s issues.

Summary of Melissa Skolfield’s Remarks

Melissa Skolfield, the Vice President for Communications at Brookings, made brief remarks adding to Leibovitz’s comments. Her primary point was that trying to maximize attention throughout the year requires a different strategy than trying to maximize attention the day the annual CWI is released. She recommended carefully considering how to reach each of the major audiences FCD wants to reach with the CWI, including the public, parents, researchers, and policymakers. Different strategies may be required for each of these groups. She suggested that FCD develop a 12-month strategy for reaching each of these groups.

Comments by Participants

1. There was debate about whether emphasizing a particular domain on the day the CWI is released adds to coverage given to the overall index or detracts from it.
2. In interpreting the meaning of the overall CWI score or the scores of any of the domains, some attention should be given to what the score means for children of different ages.
3. One suggestion for drawing more attention to the CWI was to devise a clever and appealing way to express the score or changes in the score – something that will seize people’s attention like the Doomsday Clock used by the *Bulletin of Atomic Scientists*.
4. Another idea was to break down the CWI results each year into ten or more specific stories and then approach organizations that have magazines or newsletters and try to get them to agree to take two or three stories during the course of a year.
5. Another proposal was to find some rationale to argue that the CWI should be at 120 or 130 and then publicize actions that parents, professionals, and policymakers could take to achieve the target score. A similar suggestion was to compare the U.S. with other nations as a way to stimulate support for action to improve the national score.
6. There seemed to be agreement that it would be important to do careful analysis of policies that might improve child well-being in any of the particular domains. It does little good to recommend policies or practices that might not have the hoped-for impact. One participant observed that he believes the interpretations given to the CWI in some of the annual releases have not been well-grounded in research. He cautioned that interpretations of the meaning of changes in the overall index or any of the domains should be well-supported by research.
7. Some concern was raised that the CWI did not have the scientific status of other widely-known index measures like the Consumer Price Index.