Summary record: Experts’ Meeting of the OECD Longitudinal Study of Skill Dynamics

The Board is invited to note the conclusions of the Experts’ Meeting of the OECD Longitudinal Study of Skill Dynamics (2-3 February 2015, the Hague, the Netherlands).

Koji Miyamoto [koji.miyamoto@oecd.org, tel: +33 1 4524 8554]
Maria Huerta [maria.huerta@oecd.org, tel: +33 1 4524 9612]
EDUCATION AND SOCIAL PROGRESS (ESP)

SUMMARY RECORD:

EXPERTS’ MEETING OF THE OECD LONGITUDINAL STUDY OF SKILL DYNAMICS


Conceptual Framework for Social and Emotional Skills

Introduction

1. The fourth Experts’ Meeting on Skill Instruments of the OECD Longitudinal Study of Skill Dynamics was held on 2-3 February 2015 at the Dutch Ministry of Education, Culture and Science, in The Hague. The main objective of the meeting was to discuss the conceptual framework of social and emotional skills constructs to be employed in the Longitudinal Study of Skill Dynamics. More specifically, this meeting was designed to gather recommendations by experts on:

   • The comprehensiveness of the proposed social and emotional skills constructs, including higher-order factors and lower-level facets;
   • The predictive validity and cross-cultural relevance of the proposed constructs;
   • The malleability of the proposed constructs between Grades 1 and 12;
   • The methods to assess the proposed constructs during the feasibility study.

2. To this end, the Secretariat invited a small group of experts on social and emotional skills. The list of participants can be found in Annex 1.

Experts’ key messages included:

   • The proposed conceptual framework for assessing social and emotional skills of school-aged children was considered highly appropriate. The five domains of social and emotional skills provide a useful and appropriate structural basis for the framework due to its solid empirical foundation. The extensive review of other existing frameworks was much valued, as it confirmed the high degree of convergence with the five proposed domains. However, experts suggested that additional constructs and lower-level facets should be considered and explored in the feasibility study. They include: self-beliefs and identity (sense of belonging to peers, school, community), time preference, academic self-efficacy, attitudes, motivation, interests, values and social axioms among others. A better labelling for higher-order factors (or domains) and for items at the facet level is needed in order to improve communication with diverse stakeholders.

   • Anchoring vignettes were regarded as a useful tool to adjust for measurement biases associated with rating scales. Although there is still limited evidence on the effectiveness of this technique, it was agreed that it should be used in the feasibility study for student self-reports and for student reports by their parents and teachers.

3. Rien Rouw, the Dutch CERI Governing Board member, welcomed the experts and expressed the longstanding interest and support The Netherlands has had in the ESP project. He emphasised the relevance of this work for his home country, especially for their current debate on a curriculum reform on 21st century skills. Next, Koji Miyamoto (OECD) introduced the meeting by providing an overview
of the proposed longitudinal study, highlighting the broad policy and research questions the study aims to address. Maria Huerta (OECD) then presented the non-cognitive measures used in PISA and PIAAC, and emphasised the need to align the ESP conceptual framework on social and emotional skills with the non-cognitive measures that have been assessed through other OECD activities.

**Session 1: The ESP social and emotional skills conceptual framework**

4. The first session started with a presentation by Oliver John (University of California at Berkeley) and Filip de Fruyt (Ghent University) who described the rationale of the conceptual framework and the process used to derive the new framework for social and emotional skills. They highlighted why they used the Big Five Inventory as the structural basis: 1) they are relatively uncorrelated with IQ; 2) they are universal, and less likely to be culture-specific; 3) they replicate across grades and cultures; 4) they predict important life outcomes; and, 5) they have been derived by many independent investigators. He emphasised that the Big Five Inventory could be seen as an overarching structural model that organises a number of items into one coherent taxonomy. He explained that the proposed framework builds on the Big Five as the basis and supplemented with information from other theoretical frameworks, resulting in a more comprehensive model. While being satisfied with the structure and content of the framework, he highlighted the importance of improving the labels of some domains (e.g., agreeableness and extraversion).

5. Charles Fadel (Center for Curriculum Redesign (CCR)) presented the CCR’s Character Qualities Framework by webex. The CCR’s latest framework includes six qualities: mindfulness, curiosity, courage, resilience, ethics and leadership. He highlighted the need to make the labels of constructs appealing to different stakeholders. He argued that this can be done by making them simple, crisp and explicitly linked to the education goals. He also emphasised the importance of aligning their framework with that of the OECD.

6. Patrick Kyllonen (ETS) strongly supported the proposal. He agreed on using the Big Five as the basis of the conceptual framework, because it comes with a strong scientific foundation and it has been replicated across ages and cultures. He welcomed the efforts made to look at all frameworks and see how they merged into the Big Five. He acknowledged that there is still no consensus on the most appropriate methodological approaches to measure social and emotional skills. Finally, he emphasised two additional challenges of this work: 1) addressing the strong cultural belief that personality is fixed and not malleable; and 2) finding the appropriate labels for higher-order factors and lower-level facets. These have to be carefully identified, as the choice would affect the attractiveness of this project to stakeholders.

7. Filip de Fruyt (Ghent University) argued that while Big Five constructs do not have the same importance across cultures, they still work in different countries. He also acknowledged that the current framework considers outcomes at the individual level only, but that it has the potential to also examine outcomes at the societal level: to see how people connect, collaborate, interact, which is also of interest to policymakers. After experts’ comments, he added that the framework could bring all the different proposed models together and look at the commonality.
8. The participants were invited to comment, discuss and suggest alternative options.

The experts:

- AGREED that the current proposal provides a strong conceptual framework to measure social and emotional skills across ages and countries. However, they SUGGESTED exploring constructs that are particularly relevant within (a) the socio-cultural and media contexts facing children, and (b) Asian cultures.

- More specifically, the experts SUGGESTED including constructs that may not be fully or partially covered in the proposed framework, such as self-beliefs and identity (sense of belonging to peers, school, community, the world?), connectivity (relationship with peers or community, use of social media), social axioms (social flexibility in an uncertain future, fate control), time perspective (prospection), risk aversion, attitudes, growth mindset (instead of locus of control) entrepreneurial skills, social responsibility, academic self-efficacy, interests, attitudes and motivations, communication skills, second language and religious aspects, and tactics of conflict avoidance.

- AGREED that many of the proposed constructs could be mapped onto the proposed framework. Experts SUGGESTED that constructs or facets that cannot be well mapped should be treated separately (e.g., attitudes, growth mindset, social axioms (i.e., beliefs about people, social groups, social institutions, the physical environment, or the spiritual world).

- SUGGESTED including aspects that are not only relevant now but also in the future. For example, respect, valuing diversity, and conflict resolution are key aspects for a well-functioning society and are likely to become even more relevant in a globalised world.

- AGREED that there is a need to develop better labels, so the framework is more attractive and useful for schools, teachers, parents and students.

Session 2: Predictive validity and cross-cultural relevance

9. This session discussed whether the proposed framework covers the most relevant constructs and lower-level facets across countries. Filip de Fruyt started his intervention by showing the cross-cultural relevance of the five factor model (FFM), which the proposed framework is based on. This classification model has been replicated in many cultures and has worked well, suggesting that personality traits are common across different cultural groups. Second, he explained the predictive validity of the FFM. He argued that there is a complex interplay of skills and outcomes (e.g., equifinality and multifinality). Hence, to understand how individuals develop it is necessary to look at this interplay. There are linear and curvilinear relationships, where more is not always better (e.g., industriousness) — and there may be an optimum level of skills. Third, he presented evidence on how social and emotional skills influence outcomes, such as academic performance, employability, health-related indicators and human resources competencies.

10. The three discussants, namely Michael Bond (Hong Kong Polytechnic University), Virginia Kwan (University of Arizona) and Fons van der Vijver (University of Tilberg) considered the proposed framework appropriate and comprehensive.

11. Michael Bond (Hong Kong Polytechnic University) proposed the use of social axioms (generalised beliefs about people, social groups, social institutions) to more fully capture cultural variations and account for social behaviour. He argued that beliefs about oneself and beliefs about the world are important to explain social behaviour. He suggested looking into constructs that assess true harmony,
close to the ethics category presented, and capture the individual’s capacity to develop strategies for conflict avoidance. There are instruments for assessing these items that have been tested in China and many other countries. These sets of measures may be of interest for some participating countries that can be classified as collectivistic. Michael Bond also suggested including measures of genetics; and, asking parents about their activities with children.

12. Virginia Kwan (University of Arizona) agreed that the Big Five is a comprehensive and flexible tool to study social and emotional skills across cultures. She confirmed that increasing evidence has demonstrated the universality of the Big Five taxonomy of personality. She welcomed the review in the report on culture-specific personality dimensions. She emphasised that is important to keep in mind that social values attached to personality traits vary across cultures (e.g., assertiveness is not a desirable characteristic in collectivistic cultures). She suggested considering the following lower-level facets: fear of happiness, self-efficacy, prospection, and narcissism.

13. Fons van der Vivjer (University of Tilberg) claimed that the framework is appropriate in the personality dimensions, and suggested considering additional aspects. For example, he considered that the framework should account for the impact of the context on skill development. He pointed out the following aspects as missing: personal and social identity; religion (important for identity); negative aspects of personality; and measures on global issues, such as interconnectivity (social media, second language). He suggested that the feasibility study could explore more constructs than those the main study will collect.

The experts:

- AGREED that the proposed constructs (including lower-level facets) generally reflect social and emotional characteristics that are highly appreciated across cultures and linguistic spaces. Many of them predict a range of children’s educational, labour-market and social outcomes across countries.

- AGREED that the importance of lower-level facets differs significantly across countries. Experts argued that social and emotional skills are highly contextual, and their value changes widely across cultures.

**Session 3: Malleability during Grades 1-12**

14. This session discussed the extent to which the proposed lower-level facets are malleable during Grades 1-12. Experts provided additional insights beyond what Filip de Fruyt and Oliver John presented in the background report.

15. Kimberly Schonert-Reichl (University of British Columbia) underlined the malleability of social and emotional skills – these skills can be taught. However, her experience shows that teachers and parents believe these types of skills are fixed. She argued that this project has the potential to change the cultural view of skills. She also emphasised the importance of accounting for the developmental nature of skills in the conceptual framework. Constructs will change over time and some items will be more relevant to nurture and enhance at certain grades than others. Moreover, children will understand differently the items at different stages of development, and the instruments have to enable assessment of this progression. In addition, she suggested reflecting on the definitions of growth and development, as it is not necessarily the case that more skills bring better outcomes.
16. Angela Duckworth (University of Pennsylvania) participated via webex. She outlined that adolescence is a particularly sensitive period, with emerging evidence showing that it is a key stage of development of social and emotional skills. She welcomed the fact that the study will cover Grades 1 to 12. She believes the proposed study will shed light on who changes skills more and why. She pointed out that the study should also consider selecting the right timing of data collection. Responses may change depending on the timing of the year the data is collected, for example, with more positive answers obtained at the beginning of the school year or at the beginning of the new year. Kimberly Schonert-Reichl (University of British Columbia) added that there is a cycle in the school year, where motivation decreases: from March until the end of the school year.

17. Ricardo Primi (University of Sao Francisco) suggested developing a scale to make inferences on changes in social and emotional skills between Grades 1 and 12. He suggested: a) developing hypotheses from personality/developmental psychologists about how social and emotional skills evolve; b) operationalising those hypotheses with items/markers/behaviours for selected levels along the scale; c) studying the developmental validity of the items; d) using construct maps and IRT that relates item response to the locations in the scales. He argued that defining levels of constructs should be included in the item development and tested in the feasibility study.

18. The experts:

- AGREED on the importance of accounting for the developmental nature of skills in the conceptual framework. Social and emotional skills change over time – their relevance and malleability differ across school grades. At present, there is little evidence on the trajectories of skill development.

- SUGGESTED that the conceptual framework define levels of development for each construct.

Session 4: Assessment methods

19. The aim of this session was to have a first discussion on promising assessment methodologies to measure the proposed constructs during the feasibility study. Oliver John (University of California, Berkeley) and Patrick Kyllonen (ETS) presented the challenges associated with assessing these types of skills, as well as available methods to adjust for measurement biases associated with rating scales.

20. Oliver John described some of the operational challenges. First, he suggested having measurements in more than one day. He explained that for many psychological constructs there is an effect of the day of measurement — e.g., there are days when one feels more extraverted than others. Hence, to examine how children’s skills change over time, the survey needs to account for the variation of measurement in different days. Hence, in order to have better estimates, assessments should be done more than once. Rigorous studies are doing it three times. However, Patrick Kyllonen warned that PISA has tried to have two days of testing, but countries have not allowed it. An alternative suggested option is to split the questionnaire into two days: have two sessions, repeating a subset of items to assess twice.

21. Oliver John went over the advantages and disadvantages of performance tests. They can be very expensive and are not without problems. He referred to the MSCEIT, the emotional intelligence test. Although an innovative and conceptually interesting undertaking, it has not proven the decisive fix for the biases associated with self-reported measures of emotional “competencies”. Moreover, most performance tests are very expensive to implement. He suggested the puppet interview of the University of California, Berkeley as an option for assessing the young children.
22. Another suggestion was to consider using assessments that examine situated personality — i.e., asking children how they perceive themselves at school and how they perceive themselves at home. This is useful when agreement between multiple respondents (convergent validity) tends to be low—parents and teachers have very different perspectives on children’s skills.

23. Patrick Kyllonen went over the need to use tools to adjust for response-style biases. These biases may cause problems for cross-cultural comparability and for others’ ratings, as there is a lack of differentiation (halo and horn effect). The methods to address these problems include: a) situational judgement tests; b) forced choice and, c) anchoring vignettes. However, these methods have a number of limitations. For instance, situational judgement tests take a long time to get reliable measures and are very expensive to develop and execute. PISA has tried to use these tests, but they were so complex for test takers that they ended picking children’s IQ instead of social and emotional skills. The forced choice instruments have improved cross-cultural comparability. Anchoring vignettes have also shown to work to improve cross-cultural comparability. He mentioned that there is no need to apply vignettes for all items, as one can use responses to some items to correct other scales.

24. Ricardo Primi (University of Sao Francisco) commented on the assessment methods. He agreed that vignettes are worth trying since they can help to detect and correct response styles and group reference biases. In addition, the potential application to measure growth is interesting. However, he pointed out that to conduct vertical equating it is necessary to create vignette sets that expand across grades (from grade 1 until grade 12).

25. The experts:

- AGREED that the feasibility study could include more items than those needed for the main study. They also AGREED that the study would benefit from informing schools, teachers and students what they can do with these data.

**Next steps**

Following the experts’ recommendations, the Secretariat proposed the following steps to further improve and finalise the conceptual framework:

a. Revise the conceptual framework, integrating some of the suggested constructs and reflecting on the appropriateness of the labels (higher-order factors and lower-level facets).

b. Further align our work with PISA (global competency).

c. Finalise conceptual framework after taking into account comments from the Steering Group and the CERI Governing Board. Contractors will use the conceptual framework approved by the CERI GB to explore instrument development from Fall 2015 onwards.
<table>
<thead>
<tr>
<th>First Name</th>
<th>Last Name</th>
<th>Company Name</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael H.</td>
<td>Bond</td>
<td>Hong Kong Polytechnic University</td>
<td><a href="mailto:michael.bond@polyu.edu.hk">michael.bond@polyu.edu.hk</a></td>
</tr>
<tr>
<td>Filip</td>
<td>De Fruyt</td>
<td>University of Ghent</td>
<td><a href="mailto:Filip.DeFruyt@UGent.be">Filip.DeFruyt@UGent.be</a></td>
</tr>
<tr>
<td>Andre</td>
<td>de Moor</td>
<td>Ministry of Education, Culture and Science, the Netherlands</td>
<td><a href="mailto:a.p.g.demoor@minocw.nl">a.p.g.demoor@minocw.nl</a></td>
</tr>
<tr>
<td>Angela</td>
<td>Duckworth</td>
<td>University of Pennsylvania</td>
<td><a href="mailto:angela.duckworth@gmail.com">angela.duckworth@gmail.com</a></td>
</tr>
<tr>
<td>Charles</td>
<td>Fadel</td>
<td>Center for Curriculum Redesign (CCR)</td>
<td><a href="mailto:charlesfadel@gmail.com">charlesfadel@gmail.com</a></td>
</tr>
<tr>
<td>Oliver P.</td>
<td>John</td>
<td>University of California, Berkeley</td>
<td><a href="mailto:o_johnx5@berkeley.edu">o_johnx5@berkeley.edu</a></td>
</tr>
<tr>
<td>Meesook</td>
<td>Kim</td>
<td>Korean Educational Development Institute (KEDI)</td>
<td><a href="mailto:meek@kedi.re.kr">meek@kedi.re.kr</a></td>
</tr>
<tr>
<td>Virginia</td>
<td>Kwan</td>
<td>Arizona State University</td>
<td><a href="mailto:Virginia.Kwan@asu.edu">Virginia.Kwan@asu.edu</a></td>
</tr>
<tr>
<td>Patrick</td>
<td>Kyllonen</td>
<td>Educational Testing Service (ETS)</td>
<td><a href="mailto:pkyllonen@ets.org">pkyllonen@ets.org</a></td>
</tr>
<tr>
<td>Ted</td>
<td>Reininga</td>
<td>Ministry of Education, Culture and Science, the Netherlands</td>
<td><a href="mailto:f.k.reininga@minocw.nl">f.k.reininga@minocw.nl</a></td>
</tr>
<tr>
<td>Ricardo</td>
<td>Primi</td>
<td>University of São Francisco</td>
<td><a href="mailto:Ricardo.Primi@usf.edu.br">Ricardo.Primi@usf.edu.br</a></td>
</tr>
<tr>
<td>Rien</td>
<td>Rouw</td>
<td>Ministry of Education, Culture and Science, the Netherlands</td>
<td><a href="mailto:m.a.rouw@minocw.nl">m.a.rouw@minocw.nl</a></td>
</tr>
<tr>
<td>Daniel</td>
<td>Santos</td>
<td>University of São Paulo</td>
<td><a href="mailto:daniel.ddsantos@gmail.com">daniel.ddsantos@gmail.com</a></td>
</tr>
<tr>
<td>Kimberly</td>
<td>Schonert-Reichl</td>
<td>University of British Columbia</td>
<td><a href="mailto:kimberly.schonert-reichl@ubc.ca">kimberly.schonert-reichl@ubc.ca</a></td>
</tr>
<tr>
<td>Yusuke</td>
<td>Takahashi</td>
<td>Kyoto University</td>
<td><a href="mailto:takahashi.yusuke.3n@kyoto-u.ac.jp">takahashi.yusuke.3n@kyoto-u.ac.jp</a></td>
</tr>
<tr>
<td>Yuka</td>
<td>Uzuki</td>
<td>National Institute for Educational Policy Research (NIER), Japan</td>
<td><a href="mailto:yuzuki@nier.go.jp">yuzuki@nier.go.jp</a></td>
</tr>
<tr>
<td>Fons J. R.</td>
<td>Van de Vijver</td>
<td>University of Tilburg</td>
<td><a href="mailto:fons.vandevijver@tilburguniversity.edu">fons.vandevijver@tilburguniversity.edu</a></td>
</tr>
<tr>
<td>Rolf</td>
<td>Van Der Velden</td>
<td>Research Centre for Education and the Labour Market</td>
<td><a href="mailto:r.vandervelden@maastrichtuniversity.nl">r.vandervelden@maastrichtuniversity.nl</a></td>
</tr>
<tr>
<td>Katja</td>
<td>Anger</td>
<td>OECD</td>
<td><a href="mailto:Katja.ANGER@oecd.org">Katja.ANGER@oecd.org</a></td>
</tr>
<tr>
<td>Arno</td>
<td>Engel</td>
<td>OECD</td>
<td><a href="mailto:Arno.ENGEL@oecd.org">Arno.ENGEL@oecd.org</a></td>
</tr>
<tr>
<td>Maria</td>
<td>Huerta</td>
<td>OECD</td>
<td><a href="mailto:Maria.HUERTA@oecd.org">Maria.HUERTA@oecd.org</a></td>
</tr>
<tr>
<td>Koji</td>
<td>Miyamoto</td>
<td>OECD</td>
<td><a href="mailto:Koji.MIYAMOTO@oecd.org">Koji.MIYAMOTO@oecd.org</a></td>
</tr>
</tbody>
</table>