Cognadev Technical Report Series

5

27th May, 2015

Retest Reliability of the Cognitive Process Profile (CPP)

Computed using a sample of 87 students undertaking an Accounting degree course at a South African university

Contents

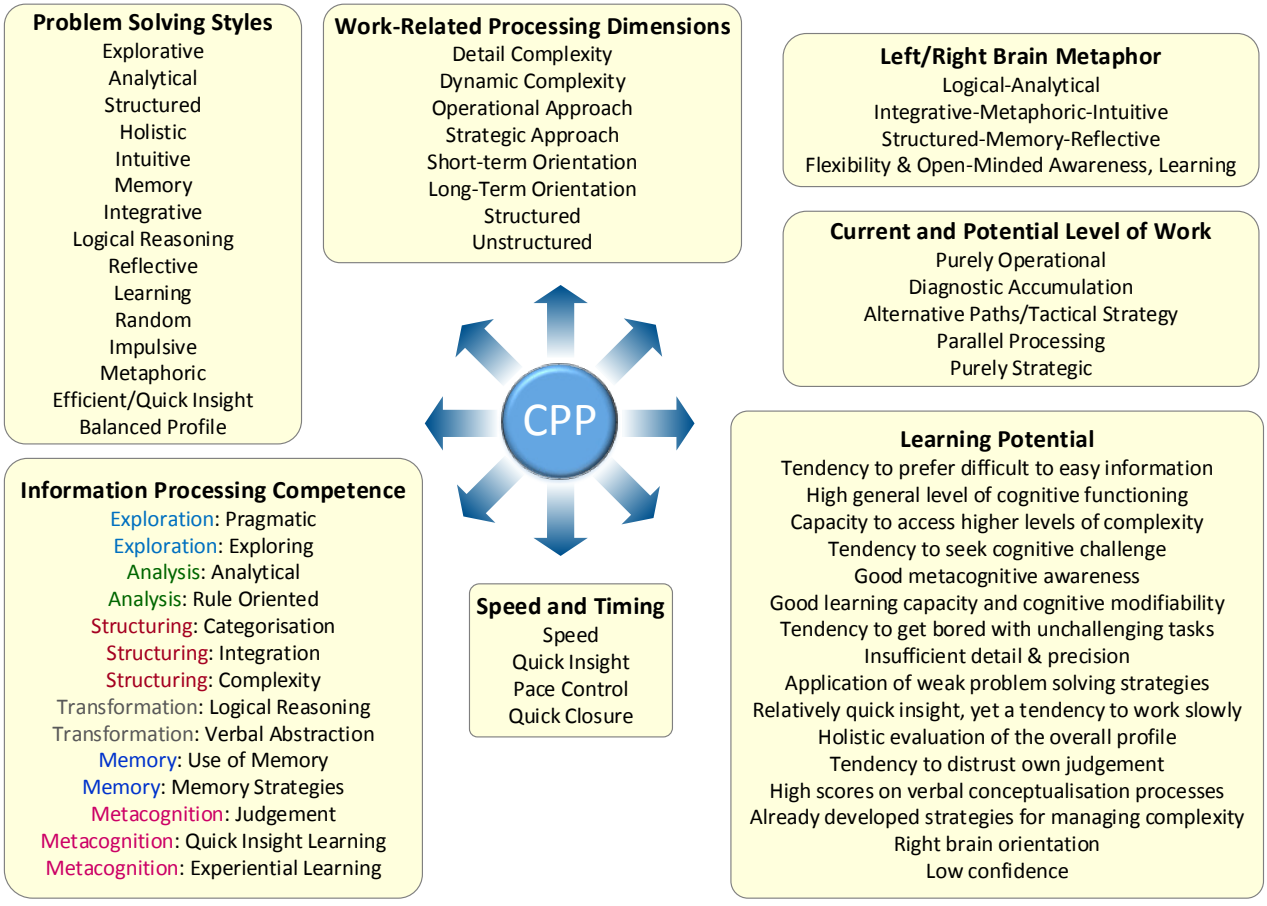
Tables

Figures

1. The Cognitive Process Profile (CPP)

The CCP is the practical application designed to assess the specific processes and concepts shown in Figure 1. It measures intellectual functioning in terms of constructs such as judgement and decision making, strategising, generalist versus specialist orientation, creativity, complexity preferences and other thinking and problem solving factors related to professional, managerial and executive functioning. It is an advanced computerised assessment technique; using simulation exercises. Subjects are monitored in terms of their preferences and capabilities in exploring, analysing, structuring, transforming, remembering, and learning information and making decisions, or exercising their judgement. The results can serve both as a source of personal understanding and development, as well as being linked to job-related performance. Figure 1 summarises the key processing components and styles.

Figure 1: The constructs reported on by the CPP



The CPP assessment consists of a task requiring the deciphering of hieroglyphic messages. It was designed to externalise and track each of the thinking processes specified in Figure 1, and their many subcomponents. While completing the test, a person explores, links, structures, transforms, remembers, clarifies and monitors his/her actions on the computer screen using a computer mouse. All the "movements" made on the computer screen are saved as the person traverses the test. At the end of each task, the person provides his/her interpretation of the symbolic message (normally a one-line statement) by keying it into the computer. A "scoring and statement parsing system" subsequently integrates all these movements and story interpretations, which are subsequently analysed using more algorithms to produce the CPP report. For the CPP to measure the various concepts detailed in the different sections of the report, the cognitive processes are grouped and analysed in many different ways. These intricate groupings often overlap, and the analysis performed by the software is highly complex.

2. Study details

2.1 Sample

87 students undertaking an Accounting degree course at a South African university comprised the sample who completed the CPP twice, along with the Learning Orientation Index (LOI) assessment. Specific gender information was not available at the time of analysis, although judging by the forenames, the majority were male students. Their ages at the first CPP assessment are provided in Table 1.

Table 1: Sample participant ages at the first assessment

	Descriptive Statistics						
Variable	Valid N	Mean	Median	Minimum	Maximum	Std.Dev.	Skewness
Age at Assessment	87	24.74	24	21	32	2.461	1.141

The descriptive statistics for the retest durations (in days) between the 1st and 2nd occasion CPP assessments are provided in Table 2. The median duration is **almost 2 months** (7.9 weeks), with the range between 7 and 9.3 weeks.

Table 2: Summary statistics for retest durations (in days)

	Descriptive Statistics						
Variable	Valid N	Mean	Median	Minimum	Maximum	Std.Dev.	Skewness
Retest Duration	87	54.89	55	49	65	2.678	0.832

2.2 Estimating retest reliability

Gower[1] agreement indices (see Appendix 1) were used throughout to express magnitude agreement, as what matters here is answering the simple question "*how closely do the retest occasion observations agree with one another?*", and not "*do observations on the first occasion possess a monotonic relationship with those on the second occasions?*"

Indexing monotonic relationship (Pearson correlation, gamma, some ICCs, Cronbach alpha) is not relevant to the assessment of retest 'reliability'; what we want to see is how closely our scores on one occasion are *reproduced* on the second. For that we need to preserve the actual magnitudes of our observations and not remove that information via a standardization transformation.

A presentation on these issues, showing empirically the inaccuracy of ICCs and Pearson coefficients (and r_{wg}[2]) for assessing retest reliability, entitled: *Interrater Reliability: measuring agreement and nothing else* can be downloaded from: http://www.pbarrett.net/issid/issid2009.html.

Relative to the maximum possible absolute (*unsigned*) discrepancy between the two pairs of observations, the Gower discrepancy coefficient indicates the % average absolute discrepancy between all pairs of observations. When expressed as a similarity coefficient (by subtracting it from 1), it indicates the % average similarity between all pairs of observations.

So, a Gower similarity coefficient of 0.90 indicates that relative to the maximum possible absolute (*unsigned*) discrepancy between them, the observations agree on average to within 90% of each other's values.

[1] Gower, J.C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27, 857-874.

[2] James, L.R., Demaree, R.G., & Wolf, G. (1984) Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69, 1, 85-98.

3. Results

3.1 CPP Ranked Styles

CPP cognitive styles refer to broad response tendencies or patterns in thinking and problem-solving behaviour. These are measured by tracking a person's responses to unfamiliar information. A person's stylistic preferences when dealing with unfamiliar information, however, also tend to be used when working with familiar information. Some personality factors are indicated here, as these are sometime evident in the way a person thinks.

A person may develop specific stylistic preferences due to personality and emotional factors, cultural values, educational exposure, learning opportunities, work experience and fields of interest. In interpreting this report, the specific combination of preferred styles provides a useful indication of certain factors in the person's developmental history.

Various descriptive categories are reported on as indications of stylistic preference, namely: Explorative, Analytical, Logical, Structured, Reflective, Reactive, Trial-and-error, Integrative, Holistic, Intuitive, Quick Insight, Learning, Metaphoric and Memory approaches. A Trial-and-error or Quick Closure style may be an indication of performance anxiety, emotional or developmental factors. It may also be a valid reflection of the person's approach to unfamiliar problem-solving. Insight can be gained from interpreting the person's particular combination of stylistic preferences. The construct of "Style" also informs the identification of a suitable work environment.

For every individual, the scores for each cognitive style are rank-ordered in their report (1 = most preferred, 14 = least-preferred).

Table 3: Retest Reliability CPP Ranked Styles {most-preferred = 1, least preferred = 14}

Ranked Style	Gower	MAD
CPP - Explorative	.81	2
CPP-Analytical	.79	3
CPP Structured	.85	2
CPP Holistic	.85	2
CPP Intuitive	.88	2
CPP Memory	.76	3
CPP Logical	.79	3
CPP Impulsive	.81	2
CPP Random	.80	3
CPP Integrative	.87	2
CPP Learning	.79	3
CPP Quick insight	.78	3
CPP Reflective	.77	3
CPP Metaphoric	.82	2

Notes: MAD = Mean Absolute deviation between the paired observations.
The MAD values are rounded integers as the possible 'score' ranges are integer ranks

The median agreement index across all 14 styles is 81%, indicating that relative to the maximum possible absolute (*unsigned*) discrepancy between them, the ranks assigned to styles across occasions agree *on average* to within 81% of each other's values.

3.2 CPP Levels of Work

The CPP links a person's cognitive profile to the cognitive requirements of specific operational and strategic work environments. Algorithms are used to compare the qualitative and quantitative characteristics of a person's profile to the requirements of five work environments. The profile qualities considered include a person's:
(a) stylistic preferences,
(b) the units of information used in processing,
(c) judgement and decision making tendencies, as well as
(d) eight job-related processing dimensions.

The work environments specified reflect the Stratified Systems Model (SST) of Jaques, the Viable Systems Model (VSM) of Beer and Prinsloo's work on cognitive complexity.

Both 'current' and 'potential' work environments are indicated but no time frames are given to predict the person's readiness to progress from the current to the potential level as this depends on many different factors including opportunity and motivation.

The CPP assigns an ordered-class 'score' to a respondent, ranging within five 'Levels of Work', for both current and potential Level of Work designations:

1. **Pure Operational**: individuals who show less interest in intellectual complexity, vagueness and cognitive challenge.
2. **Diagnostic**: can be quite analytical, but still show a need for structure in the form of technical guidelines and/or previous experience.
3. **Tactical Strategy**: no longer rely on linear processing, but prefer viewing issues in terms of tangible systems and the interaction between observable system elements.
4. **Parallel Processing**: those with the capacity to accommodate novelty, vagueness, dissonance and fragmentation, all of which require the cognitive skills of integration and innovation.
5. **Pure Strategic**: functioning is characterised by a strong Intuitive and Holistic "big picture" inclination.

Table 4 reports the agreement indices for the current (cLOW) and potential (pLOW) Levels of Work attributes.

Table 4: Retest Reliability CPP Levels of Work {cLOW range 1-4, pLOW range 1-5}

Level of Work (LOW)	Gower	MAD
CPP Current LOW	.87	.40
CPP Potential LOW	.87	.53

Note: MAD = Mean Absolute deviation between the paired class-categories

Indicating that relative to the maximum possible absolute (*unsigned*) discrepancy between them, the ordered class-categories agree *on average* to within 87% of each other's values.

It is important to view these data in more detail, as the summary agreement index doesn't really convey where discrepancies are occurring between occasions. Tables 5 and 6 report the occasion-1 vs occasion-2 cross-tab frequencies for Current and Potential Levels of Work attributes.

Table 5: CPP Current Level of Work cross-tab frequencies

Summary Frequency Table : CPP Current Level of Work, 2-occasions Marked cells have counts > 10 (Marginal summaries are not marked)					
CPP-cLOW-1	CPP-cLOW-2 1	CPP-cLOW-2 2	CPP-cLOW-2 3	CPP-cLOW-2 4	Row Totals
1	17	12	0	0	29
2	2	32	8	3	45
3	0	3	4	3	10
4	0	0	1	2	3
All Grps	19	47	13	8	87

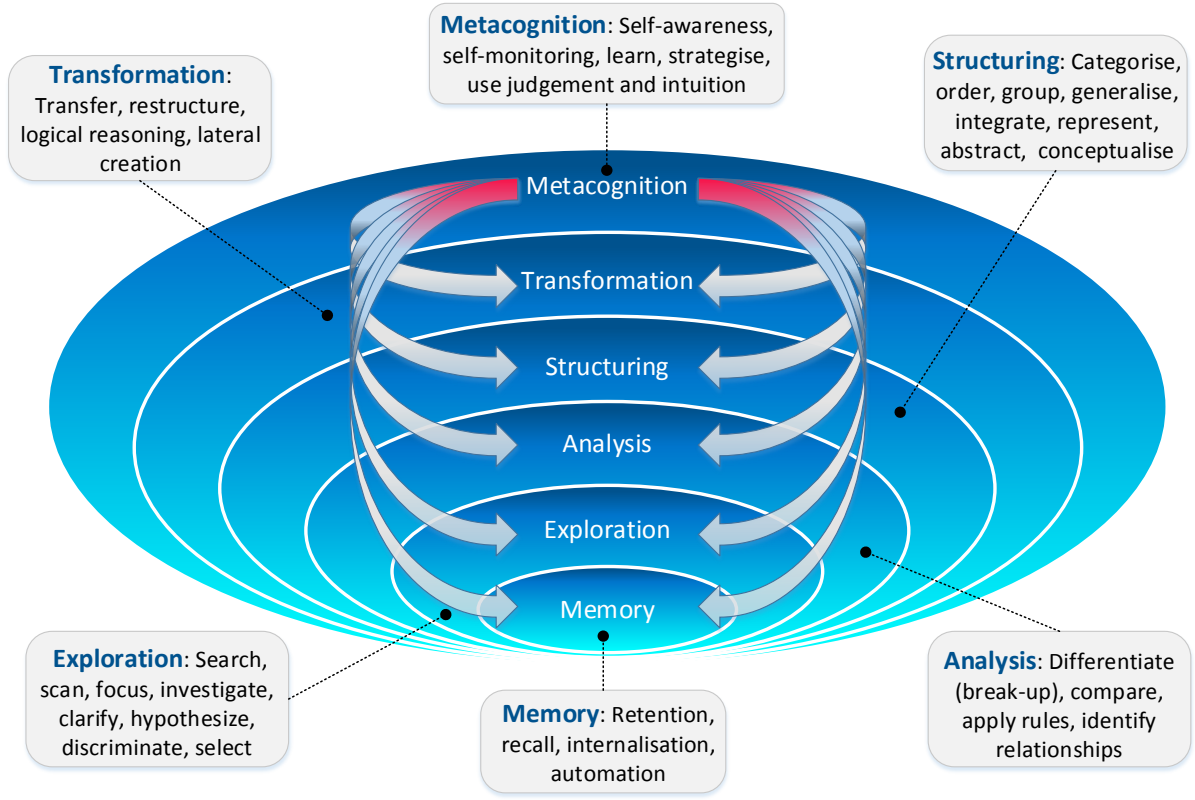
Table 6: CPP Potential Level of Work cross-tab frequencies

Summary Frequency Table : CPP Potential Level of Work, 2-occasions Marked cells have counts > 10 (Marginal summaries are not marked)						
CPP-pLOW-1	CPP-pLOW-2 1	CPP-pLOW-2 2	CPP-pLOW-2 3	CPP-pLOW-2 4	CPP-pLOW-2 5	Row Totals
1	1	4	1	0	0	6
2	1	26	15	3	0	45
3	0	7	9	5	1	22
4	0	0	3	10	1	14
All Grps	2	37	28	18	2	87

3.3 CPP Process Scores

These index cognitive functioning as a classification into six major processing categories. These functional processing categories can be represented as a *holon* where each successive process includes and transcends the previous one(s). Figure 2 summarises the processing categories and their subdivisions.

Figure 2: The cognitive processes assessed by the CPP



The dynamic functioning of the processes are explained by the theoretical processing model on which the CPP assessment is based. Figure 3 provides the brief descriptions of the processing categories and their subdivisions.

Figure 3: The 14 CPP processing competencies

Processing Competency		Description
Memory	**Use of Memory**	A tendency to rely on memory and to concentrate on the task
	Memory Strategies	Effectiveness of memory strategies
Exploration	**Pragmatic**	Practical orientation (asking whether things will work in practice). Determining relevance in structured contexts
	Exploration	The effectiveness, depth and width of exploration
Analysis	**Analysis**	Working systematically. Detailed and precise in differentiating between, and linking, elements
	Rules	A focus on rules
Structuring / Integration	**Categorisation**	Creating external order, categories and reminders. Structuring tangibles
	Integration	Synthesis of ambiguous / discrepant / fragmented information
	Complexity	The preferred level of complexity and the unit of information used
Transformation	**Logical Reasoning**	The disciplined, logical following through of reasoning processes
	Verbal Conceptualisation	Unusual / flowery / creative and/or abstract verbalisation and conceptualisation
Metacognition	**Judgment**	Capitalising on intuitive insights to clarify unstructured and vague information
	Quick Insight Learning	A tendency to grasp new concepts and acquire knowledge and understanding relatively quickly
	Gradual Improvement Learning	A preference for practical or experiential learning

Within this retest dataset, all 14 process scores were expressed on a 1 to 7 scale (low to high). Table 7 provides the Gower agreement indices between the two-occasion CPP assessments.

Table 7: Retest Reliability CPP Process Scores {score range 1-7}

Process Score	Gower	MAD
CPP-Pragmatic	.83	1.0
CPP-Exploration	.88	0.7
CPP-Analytical	.79	1.3
CPP-Rule-Oriented	.86	0.9
CPP-Categorisation	.84	1.0
CPP-Integration	.88	0.7
CPP-Complexity	.85	0.9
CPP-Logical Reasoning	.79	1.3
CPP-Verbal Conceptualisation	.77	1.4
CPP-Use of Memory	.84	1.0
CPP-Memory Strategies	.86	0.9
CPP-Judgement	.85	0.9
CPP-Quick Insight Learning	.87	0.8
CPP-Gradual Improvement Learning	.79	1.3

Note: MAD = Mean Absolute deviation between the paired class-categories

The median agreement index across all 14 processes is 85%, indicating that relative to the maximum possible absolute (*unsigned*) discrepancy between them, the ranks assigned to processes across occasions agree *on average* to within 85% of each other's values.

The mean absolute deviation is 1.0 (rounded from 0.95). That is, each process score lies with ±1 of each other's values – bearing in mind the range of the scores is between 1 and 7.

3.4 Conclusion

Overall, the CPP ranked styles, Levels of Work class-categories, and processing scores showed good retest reliability over a 2-month duration; compatible with and in most cases exceeding such reliabilities found using conventional psychometric tests over this duration.

4. The Logic of CPP retest assessment: An Advisory

Because the CPP capitalizes on a person's cognitive response to new and unfamiliar information, the first CPP is always the most valid – particularly if the person's performance has not been affected by extreme performance anxiety or demotivation. Note that a manageable degree of performance anxiety may even improve concentration.

CPP re-assessment, especially where the first CPP can be regarded as valid, should therefore be postponed by at least 4 to 5 years or more, if possible.

However, at times it is useful to evaluate the impact of developmental initiatives, work exposure, maturity, changes in attitude and interest on cognition, or to reassess those with invalid reports. The second set of CPP results then has to be interpreted qualitatively. Cognadev can assist consultants in doing so.

Higher CPP scores are often obtained with the second assessment, especially in the case of operational profiles. Certain processing dimensions also tend to improve with a second assessment, such as the Analytical skills. Other dimensions are more resistant to change. These include the Potential level of work indication, the Units of information or Complexity preferences, Integration and Judgement skills. In the case of Strategic profiles, the second set of results may, however, be somewhat lower than the first, as the candidate is likely to approach a familiar problem somewhat differently from an unfamiliar problem. This may include taking short cuts based on what is already known and capitalizing on memory.

Seeing that the CPP capitalises on measuring thinking processes in response to unfamiliar problems, the most valid results are obtained with a first assessment. Test administrators should therefore take care that test candidates are calm, rested and motivated to complete their first CPP assessments. The results may remain a valid reflection of the person's approach for a long period which exceeds the 5 year cut-off point specified for a valid reassessment. However, developmental initiatives, personal motivation and work experience may improve a person's cognitive approach and this can be assessed for by means of the CPP after a certain period of time.

Appendix 1: The Gower Agreement Coefficient

Relative to the maximum possible absolute (*unsigned*) discrepancy between the two pairs of observations, the Gower *discrepancy* coefficient indicates the % average absolute discrepancy between all pairs of observations. When expressed as a similarity coefficient (by subtracting it from 1), it indicates the % average similarity between all pairs of observations. The Gower coefficient varies between 0 and 1 (or 0% and 100%).

So, a Gower similarity coefficient of say 0.90 indicates that relative to the maximum possible absolute (*unsigned*) discrepancy between them, the observations agree on average to within 90% of each other's values.

If you change the value of that maximum possible discrepancy, then the Gower coefficient will change to reflect this, as the discrepancies between pairs of observations are divided (scaled) by that maximum possible discrepancy value. E.g. if two observations differ by 5, and the measurement range of each observation is 10, then the relative discrepancy is 0.5. However, if the measurement range for each observation was say 100, then the relative discrepancy would be just 0.1.

But that's the whole point of the Gower, it tells you how discrepant (or similar) observations are, RELATIVE to how maximally discrepant they could have been.
A 5-point difference in a 10-point maximum measurement range is substantial.
A 5-point difference between observations within a 100-point measurement range is trivial.

The equation for the Gower similarity index is:

$$Gower_{similarity} = 1 - \left[\frac{\sum_{i=1}^{n} \left(\frac{|obs_{1i} - obs_{2i}|}{range} \right)}{n} \right]$$

$n =$ the number of cases

$range =$ the maximum possible discrepancy between the two attribute/variable magnitudes (100-0)

$obs_{1i} =$ the observed value for case i of n on the first occasion

$obs_{2i} =$ the observed value for case i of n on the second occasion

A free-to-download computer program for computing the Gower, along with a free bootstrap program to compute its statistical significance *(in terms of the likelihood of observing a coefficient as large as computed by chance alone)* are available from:

http://www.pbarrett.net/Gower/Gower.html and http://www.pbarrett.net/Bootstrap/Bootstrap.html