

Appendix: Technical Manual (Learning Agility)

The following appendix documents recent changes and improvements made on the Assesio Learning Agility Assessment.

Norm group update: Global norm

To reflect the current state of normal, work-related behavior, the norm group was updated based on quality standards derived from various international standards, including EFPA, COTAN, and ITC guidelines. In short, these guidelines set out criteria for various aspects of the norm group:

- **Update:** When was the norm group last updated?
- **Sample size:** How large is the norm group? Is it sufficiently large to ensure representativeness?
- **Composition:** How is the norm group composed with respect to different demographics?
- **Subgroup differences:** Are group differences sufficiently small to prevent adverse impact?

Update

Over time, what is considered normal behavior changes. Major events and crises have an impact on the way people in general behave and new generations may also challenge the existing standards. Therefore, with respect to assessments, it is highly important to update norm groups at a regular basis and make sure that all candidates and people assessed are evaluated with a norm group representing the current state and what is currently considered normal behavior, since that will provide the most valid assessment. In addition, updating the norm group keeps scores balanced and avoid too many candidates getting either high or low scores. In other words, norm updates allow for better differentiation of candidates, which in turn leads to better recruitment decisions.

According to EFPA and COTAN guidelines, a norm of the highest quality should not be older than 10 or 15 years, respectively. At Assesio, however, we are committed to checking if updates are needed at least every 2 years and update our norm groups frequently.

The new global norm replaces the 2019 norm and is based on data collected in a high-stake (selection) setting from 2020-2023.

Sample size

A good norm group consists of many people, as a high number provides greater representation and statistical certainty. The prevailing view is that the larger the sample, the better the norm group. While that is true, it very much depends on sampling procedures as well as composition with respect to different demographic characteristics. In general, norm groups that are too small run the risk of underrepresentation (e.g., too few people with a certain occupation or education level), whereas too large norm groups risk overrepresentation (e.g., too many people of a certain age or nationality).

According to EFPA, a sample size of at least 1,000 constitutes an excellent norm group (in some cases, smaller norm groups may also be sufficient depending on composition, target groups, and intended applications).

The norm group consists of 1,152 people who were selected through stratified randomization from a total of 4,491 people aged 18-70 who completed the MAP and Match-V assessments in a high-stake

setting. Statistical analyses confirmed that the norm group does not represent a biased sample, as score differences between different samples were small or negligible across scales (average Cohen's $d = 0.23$).

Composition

To ensure that a norm group is representative of all target groups and is appropriate for all intended applications, key demographic characteristics must be carefully weighted and balanced, especially those that can lead to potential score differences between subgroups.

To construct a proper global norm, the sample was stratified such that no nationality exceeded 10 % of the total norm group (the 10 % limit was chosen to minimize loss of data and ensure a proper sample size). Next, the sample was balanced on main genders (female and male), and other genders were included as well with the aim of having a representation of roughly 1 % (please note that as a result of the balancing at the last step, a few nationalities exceeded the 10 % limit by an insignificant amount). As the final sample comprised a proper range of age groups, education levels and occupations (job families), and there were no major score differences, the sample was not further stratified for any of these demographic variables.

The demographic composition of the final norm group is listed below in table A1.

Table A1. Demographic composition of the global norm for Learning Agility,

Learning Agility: Global norm	
Last updated	2024
Data collection	2020-2023
Sample size	1,152
Composition	
Purpose(s)	Selection
Age	19-65 (M = 38.7, SD = 9.32)
Gender	49.6% females 49.6% males .8% other
Nationalities	84, Max. = 11.2%
Education level (%)	
Elementary school	.3
Less than 3 years of post-secondary education	8.3
Middle/Junior high/High school	6.6
Other	4.5
PhD	3.8
Unknown	.5
3 or more years of post-secondary education	76.0
Job family (%)	
Architecture and engineering	7.6
Arts design entertainment sports and media	2.3
Building and grounds	.6
Business and financial operations	14.1

Cleaning and maintenance	-
Community and social service	.5
Computer and mathematical	3.1
Construction and extraction	1.2
Education training and library	2.3
Farming fishing and forestry	.1
Food preparation and serving related	.7
Healthcare practitioners and technical	1.9
Healthcare support	1.0
Installation maintenance and repair	3.0
Legal	1.2
Life physical and social science	1.5
Management	12.9
Military specific	.3
Office and administrative support	3.5
Other	22.4
Personal care and service	.3
Unknown	.9
Production	4.6
Protective service	.2
Sales and related	12.5
Transportation and material moving	1.4

Subgroup differences

Finally, to further substantiate the quality of the norm group composition, subgroup differences were examined with independent samples t-tests for gender (male/female) and age (dichotomized using 40 as the cut-off value) listed below in Tables A2 and A3, respectively.

Table A2. Analyses of subgroup differences for gender (male/female).

Scale	Male			Female			Comparison			
	N	M	SD	N	M	SD	Dif.	t	p	d
Change Agility	571	29.8	4.09	571	28.5	4.20	1.4	5.59	< .001	0.33
Mental Agility	571	31.8	3.70	571	30.8	3.73	1.0	4.57	< .001	0.27
People Agility	571	32.0	3.28	571	32.5	3.28	-0.5	-2.54	.01	0.15
Results Agility	571	33.0	3.99	571	32.9	3.55	0.2	0.80	.42	0.05
Self-awareness	571	32.8	3.56	571	32.3	3.61	0.5	2.55	.01	0.15
Total LA	571	159.6	13.68	571	157.0	13.70	2.6	3.21	.001	0.19

Table A3. Analyses of subgroup differences for age (above/below 40).

Scale	40+			< 40			Comparison			
	N	M	SD	N	M	SD	Dif.	t	p	d
Change Agility	479	28.4	4.19	673	29.7	4.14	-1.3	-5.39	< .001	0.32
Mental Agility	479	30.6	3.68	673	31.9	3.73	-1.3	-5.78	< .001	0.35
People Agility	479	32.2	3.08	673	32.3	3.44	-0.1	-0.35	.73	0.02
Results Agility	479	32.6	3.70	673	33.2	3.84	-0.6	-2.60	.01	0.16
Self-awareness	479	31.8	3.58	673	33.1	3.50	-1.3	-6.32	< .001	0.38
Total LA	479	155.5	13.13	673	160.2	13.93	-4.6	-5.67	< .001	0.34

These analyses showed that across scales and demographic variables, effect sizes (Cohen's d) are only small (< 0.50) or negligible (< 0.20), ranging from 0.05-0.33 for gender and 0.02-0.38 for age. Furthermore, adverse impact simulations indicated that neither of the scales show substantially different selections rates for different gender and age groups when applying moderate or lenient selection ratios. However, please keep in mind that these simulations are based on the rather theoretical assumption that people are selected based on a single scale score (which is usually not the case in actual practice). Also, Assessio does not recommend basing recruitment decisions on any single score but rather to combine scores and information derived from different sources.

In conclusion, the global norm for Learning Agility constitutes a large, well-composed sample that is suitable for the intended target groups and applications.

Scale revision: Self-awareness

As part of the continuous development of the assessment, the Learning Agility scales were revised based on the results of the preliminary validity studies cited in the technical manual. Specifically, the Self-awareness scale showed limited convergent validity with other measures of the same construct (due to a different operationalization) and a moderate internal consistency (in part due to the limited number of items used in the scale).

To construct a new scale more closely aligned with the theoretical framework, the initial pool of items to select from was expanded based on previous research. Items were then rated independently by two subject matter experts (SMEs) for face and content validity reaching a final pool of 15 items, from which the scale was constructed. These items were then examined for five key psychometric properties: Item differentiation, internal consistency, unidimensionality, local independence, and item invariance (no Differential Item Functioning). The final scale comprised a total of 10 items with acceptable internal consistency ($\alpha = .70$) and excellent levels of skewness and kurtosis (.08 and -.22, respectively).

As stated in the technical manual, the possible moderating influence of Self-awareness on the remaining LA dimensions were incorporated by weighting it into the four Agility dimensions. However, this was removed in the final scoring model as no evidence nor data was found to support this implementation.

For the remaining scales, minor adjustments were made through a similar process. These adjustments did not significantly alter the content, definition, interpretation, nor psychometric properties (including group differences and adverse impact) for any of the scales.

Additional statistical analyses

To further assess the psychometric quality of the revised scales, the intercorrelations of scale scores and their factor structure were also examined. The revised correlation matrix is shown below in Table A4 (alpha coefficients are displayed in parentheses on the diagonal).

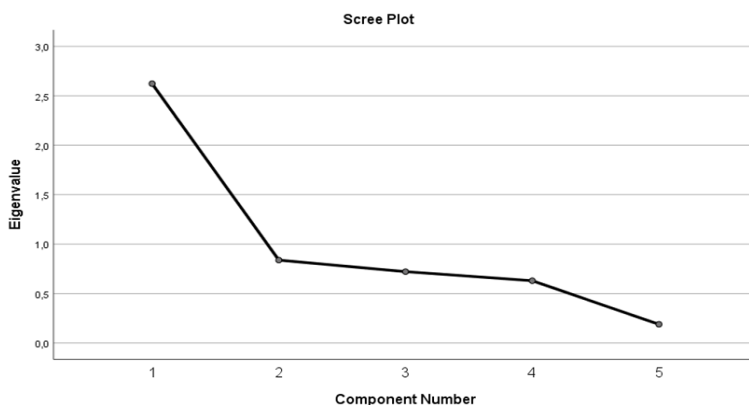
Table A4: Correlation matrix for revised Learning agility scales (alpha values in parentheses),

Scale	1	2	3	4	5	Fac. Load.
1, Change Agility	(.77)	-	-	-	-	.66
2, Mental Agility	.38	(.75)	-	-	-	.82
3, People Agility	.25	.31	(.70)	-	-	.59
4, Results Agility	.33	.29	.33	(.79)	-	.64
5, Self-awareness	.45	.80	.36	.43	(.70)	.88
6, Total LA	.69	.77	.61	.67	.84	-

As shown in the table, all subscales have moderate correlations with each other and strong correlations with total Learning Agility. This suggests that the composite scale captures all of the LA dimensions. Across scales, Self-awareness has the strongest correlation with total LA, thus emphasizing – in accordance with the theoretical framework – its key importance for Learning Agility as a foundation for learning and personal development.

The key structure of Learning Agility is further supported by a higher-order exploratory factor analysis suggesting a one-factor solution across different extraction criteria (Scree plot, Kaiser's criterion, Parallel Analysis) explaining 52.5 % of the total variance (Scree plot is shown below in Figure A1). In addition, all scales have strong factor loadings (> .40) with the overall factor as shown in the last column of table A4.

Figure A1. Scree plot for higher-order factor analysis of LA scales.



In sum, the revised scales show excellent psychometric properties in terms of content validity, internal consistency, scale correlations, and factor structure supporting the notion of Learning Agility as a unitary construct.