Psychometric Properties and Analysis of the *AQ PROFILE®*

Online Version 8.1

2009

Jerilee Grandy, PhD
Psychometric Research Consultant

PEAK LEARNING

Introduction to the AQ Profile®

The AQ Profile® (AQP)[1] is the most widely adopted method in the world for gauging human resilience. Industry-leading organizations worldwide use the AQP to screen applicants and to develop leaders, individuals, and teams. AQ is used to enhance resilience, mindset, performance, innovation, entrepreneurship, decision making, problem solving, energy, engagement, health, optimism, profitability, stock price, and competitive strength.

Harvard Business School incorporates the AQP, AQ theory and AQ methods into its prestigious executive development and MBA programs.

The standard, electronic form of the AQP is an online, interactive questionnaire designed to measure an individual's pattern of response to a broad range of adverse situations. It takes most respondents 7-10 minutes to complete the AQP 8.1.

The AQP comes in two forms. The Applicant Screening AQ Profile® is used to screen AQ among job applicants. The Developmental AQ Profile® is used for growing one's AQ. While the features that customize the AQP to these two applications differ, the central instrument, as analyzed and reported in this document, remains the same.

Purpose of This Manual

For leaders, professionals, and researchers seeking greater insight into the reliability, validity, and overall construction of the AQP, this manual provides an essential foundation. It serves as a technical supplement to the existing literature on AQ, including three bestselling books, Stoltz (1997, 2000), and Stoltz & Weihenmayer (2008).

Background

Research on the related subjects of resiliency, hardiness, optimism, psychoneuroimmunology, neuropsychology, neuroanthropology, locus of control, neurobiology, neuroeconomics, attribution theory, self-efficacy, learned helplessness and the overarching field of positive psychology, suggest that one's mindset, capacity, performance, energy, health, innovation, agility, happiness, success and longevity in life are largely determined by how one responds to adversity (Abramson, Seligman, & Teasdale, 1978; Aldwin & Gilman, 2004; Compton, 2005; Hiroto & Seligman, 1975; Irwin & Vedhara, 2005; Kobasa, 1979; Kobasa, Maddi, & Kahn, 1982; Ong, Bergeman, Bisconti, & Wallace, 2006; Peterson, Maier, & Seligman, 1993; Peterson & Seligman, 2004; Reivich & Shatte, 2002; Rotter, 1966; Seligman, 1991; Tugade, Frederickson, & Barrett, 2004; Wortman & Brehm, 1975). This research is derived from a broad range of scientific fields including cognitive psychology, health sciences, and neurology. Stoltz (1997, 2000) discusses the contributions of research in these fields to the science of AQ and to the development and use of the *AQ Profile*.

---

[1] The AQP 8.1 is derived from the paper-and-pencil form of the AQP (PEAK Learning Inc., 2008) and reflects the ongoing improvements and evolution gained over preceding versions, since 1993. To date, roughly 500,000 individuals have completed some version of the AQ Profile®.

Note: At the end of the report is a summary of terms used in this report—terminology that may be unfamiliar to the reader. Each of these terms will also be defined as it is introduced in the report.

In 2000, PEAK Learning produced a manual for interested users of the *AQP*. It was based on data from a sample of 2,414 people who took the paper-and-pencil form of the *AQP*. The manual reported norms by gender and ethnic group, reliability, and evidence for convergent and discriminant validity (Grandy, 2000). At that time, the *AQP* was called the *Adversity Response Profile (ARP)*.

Norms on the *AQ Profile*

The AQP is regularly analyzed and updated. Statistical findings occasionally alter the means and ranges for both AQ and CORE. The statistics reported here are based on a diverse sample of 1,743 employees of two global companies representing 26 countries in six regions of the world. The sample represented a broad range of job levels and responsibilities.

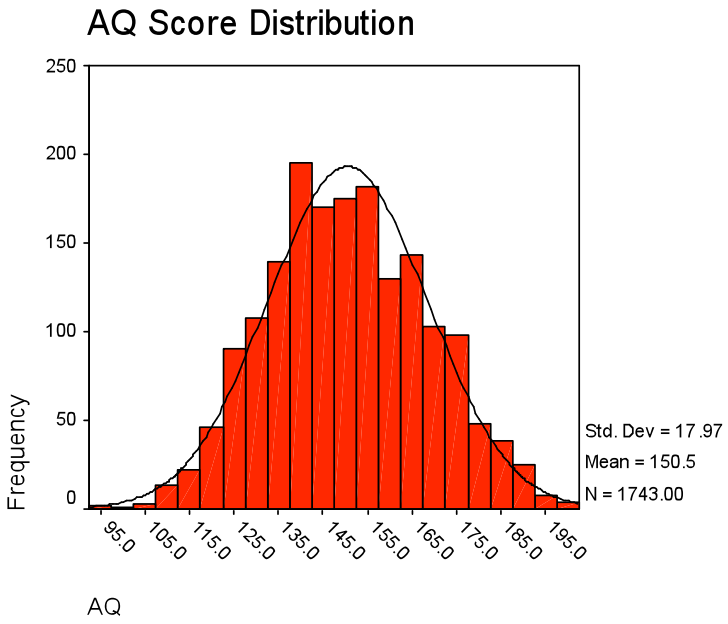
The distribution of AQ scores provides norms with which anyone taking the *AQP* can compare with his or her score. Scores on each scale of the *AQP* can range from 10 to 50, and AQ scores can range from 40 to 200. Table 1 shows the mean, standard deviation, and minimum and maximum scores on each scale and on total AQ.

Table 1. Scale Score Statistics *(N = 1743)*

	Scale				
	C	O	R	E	AQ
Mean	39.8	42.0	32.1	36.6	150.5
Std dev	6.03	6.59	5.98	5.47	17.97
Min.	15	11	11	18	96
Max.	50	50	50	50	200

The *percentile ranking* is the percentage of people scoring at or below a specified value. Table 2 shows the score associated with every fifth percentile. Half of the sample obtained AQ scores of 150 or less. Five percent obtained scores of 122 or less, and 5% obtained scores over 181. One quarter of the sample scored 138 or less, and one quarter scored over 164. The graph shows that the distribution of AQ scores is very nearly normal.

Table 2. Scale Scores Corresponding to Selected Percentile Rankings (N=1743)					
Percentile Ranking	Control	Ownership	Reach	Endurance	AQ
5	29	29	23	28	122
10	32	33	25	30	127
15	34	35	26	31	131
20	35	37	27	32	135
25	36	39	28	33	138
30	37	40	29	34	140
35	38	41	30	34	142
40	39	41	30	35	145
45	39	42	31	36	147
50	40	43	32	37	150
55	41	44	32	37	152
60	42	45	33	38	155
65	42	46	34	39	157
70	43	46	35	39	160
75	44	47	36	40	164
80	45	48	37	41	166
85	46	49	39	42	170
90	48	49	40	44	175
95	50	50	43	46	181



Relationships of AQ with Demographics

Of the 1,743 respondents, only one omitted the question on gender. Of those who answered the gender question, 56% were female. There were very small but statistically significant gender differences on all scales except Ownership. The means were one-fifth or less of a standard deviation lower for women than for men. That small a difference (called an effect size) is too

small to be regarded as meaningful[2]. The gender difference in the AQ score was extremely small (0.07 standard deviation). See Table 3.

Table 3. Scale Score Statistics by Gender						
	Female (N = 970)		Male (N = 772)			
Scale	Mean	Std. Dev.	Mean	Std. Dev.	Mean Dif.	Effect Size
Control	39.3	6.25	40.5	5.68	1.2**	0.14
Ownership	41.9	6.73	42.3	6.40	0.4	0.04
Reach	31.4	5.84	33.1	6.02	1.7**	0.21
Endurance	36.0	5.41	37.3	5.48	1.3**	0.17
AQ	148.5	17.98	153.1	17.63	1.8**	0.07

***P < 0.01*

AQ scores were found to be correlated to a small, but statistically significant, degree with age. The correlation was positive, meaning that older employees scored slightly higher on AQ and three of its four subscores. See Table 4.

Table 4. Correlations between AQ Scale Scores and Age *(N = 1125)*					
	Scale				
	C	O	R	E	AQ
Correlation	0.110**	0.125**	0.128**	0.034	0.135**

***P < 0.01*

Reliability of the *AQ Profile*

Reliability has a number of different meanings. Essentially, it refers to the consistency with which something is measured. For the *AQP*, reliability may refer to internal consistency, that is, the consistency of answers to all questions within a scale, or it may refer to the consistency of answers at two different points in time when no change in AQ has occurred during that time interval. The first of these meanings—internal consistency—is most appropriate for estimating the reliability of the *AQP* because life experiences may cause a person's AQ to rise or fall over time.

Reliability coefficients may range from 0 to 1, 1 being the highest and strongest score that can be achieved. Specifically, a reliability of 0 means that answers to questions are entirely unrelated to one another, often because they measure different traits. A reliability of 1 would mean that all answers are perfectly intercorrelated (a condition that would happen if all questions were identical or nearly identical). Realistically, a test is regarded as having "very good" reliability if its reliability coefficient is greater than roughly 0.8. Subscores, because they

[2]The largest gender difference was in the Reach score, and it was only one-fifth of a standard deviation, which is generally regarded as "small." See Cohen (1988).

are based on fewer numbers of questions, generally have lower reliabilities than do total scores. A subscore reliability greater than about 0.7 may be regarded as “very good.”

The AQ score and all four subscores were found to have high reliabilities. Table 5 shows Cronbach's coefficient alpha—a measure of the internal-consistency reliability of each scale score.

Table 5. Reliability Estimates (alpha)	
Scale	alpha
Control	0.82
Ownership	0.83
Reach	0.84
Endurance	0.80
AQ	0.91

Validity of the *AQ Profile*

Validity has two components. First, a test or questionnaire is said to be valid if it measures what it is designed, intended, and used to measure. This is called **convergent validity**. The second component is called **discriminant validity**. A test or questionnaire has discriminant validity if it does not measure traits, knowledge, or skills other than the ones it is designed to measure. Two different scales on a test, for example, should measure different things if they have different names. Sometimes the two things that are being measured are related, but they should not be identical, otherwise, there is no justification for having two scales that purport to measure two different things.

Discriminant Validity of the AQ Profile Scale Scores. To justify having four subscores, the intercorrelations among those scores should be less than their corresponding reliabilities (Campbell, 1960; Campbell & Fiske, 1959). Table 6 shows the intercorrelations of the *AQ Profile* scale scores.

Table 6. Intercorrelations of Scale Scores (N = 1743)					
	C	O	R	E	AQ
C	1.000				
O	0.494	1.000			
R	0.313	0.275	1.000		
E	0.349	0.323	0.724	1.000	
AQ	0.727	0.723	0.760	0.781	1.000

The highest correlation between scale scores is 0.724 between Reach and Endurance. The other combinations of scale scores have moderate intercorrelations. None of the

intercorrelations among scale scores is as high as the scale reliabilities, though the correlation between R and E is high enough to suggest that the two scales are measuring related but different constructs. If a person tends to generalize adversity across situations (Reach), he or she may also tend to generalize adversity over time (Endurance). Still, each of these scales shows some unique variance, so the scales are not redundant. The four scales can, therefore, be said to have demonstrated good discriminant validity. As intended, they measure different, but highly related, aspects of AQ.

Convergent Validity of the AQP. Convergent validity is the most difficult, and the most important, type of validity to establish. It is generally easy to show what a scale does not measure (discriminant validity), but to prove that any test measures, or predicts, some kind of behavior, is always a challenge. Generally, validity studies involve the correlation of test scores with some *criterion* measure, such as sales, supervisor ratings, promotions, or graduation rates. A high correlation, or statistically significant correlation, points to evidence that the test measures, or predicts, the criterion. Similarly, if individuals scoring high on the test outperform individuals scoring low on the test, we have good evidence of the validity of the test.

Validity Study

We conducted an extensive validity study on a sample of 1,130 employees of one global company. AQ theory would predict that people who are high in AQ would also perceive themselves as healthier, taking fewer prescription medicines, feeling fitter, more energetic, happier, more optimistic, successful, luckier, engaged in more exercise, experiencing less stress, and being more satisfied with their jobs. They would also be expected to take fewer sick days from work, a factor that is of considerable importance to employers.

Method. To test whether employee perceptions of their health were related to AQ, a 29-item questionnaire was designed and included with the AQ survey. The questions covered the subject's perception of specific personal health factors (digestive system, respiratory system, cardiovascular system, muscular-skeletal system, stress, fitness, energy, diet, and general health), use of prescribed medicines, exercise, perceptions of happiness, optimism, hardship, luck, success in life, and job satisfaction.

Generally there were two questions, with response options on a 7-point scale, measuring each perception. For example, one question read, "In my current job, on most days I feel…..
(1) completely involved >>>> (7) completely uninvolved." Another question read, "On most days I find my job….. (1) very enjoyable >>>>> (7) something I dread." Although both of these questions measure job satisfaction, they are directed at somewhat different aspects of satisfaction, and could be combined to form a score that is more reliable than one single question would be.

Among the additional data collected were the number of days each employee was absent over the previous year and the number of occurrences of absenteeism. If an employee missed 10 days altogether, for example, and if those days broke down into 3 days at one time, 2 days the next time, and 5 days at another time, we counted this as 3 occurrences of absenteeism.

Analyses. The analyses consisted of:

- correlations between AQ and responses to the 29 health survey questions;
- correlations between AQ and the number days absent and number of absentee occurrences over the past year;
- comparisons of the mean number of days absent and mean number of absentee occurrences for employees obtaining high AQ scores versus those obtaining low AQ scores.

Results. All 29 of the health study questions correlated significantly with AQ. When paired into the scales mentioned above – digestive system, diet, job satisfaction, etcetera – the correlations with AQ ranged from 0.104 for the cardiovascular system to 0.433 for optimism. A single question asking the number of days they participated in vigorous exercise correlated 0.111 with AQ. If there had been one more write-in question similar to the question on exercise, they could have been combined, and the resulting correlation would have been even higher.

AQ scores were correlated significantly both with days absent and number of absentee occurrences. The higher the AQ, the fewer were the number of days absent or the number of occurrences.

Correlation of AQ with days absent = -0.116, $p < 0.01$
Correlation of AQ with number of absentee occurrences = -0.182, $p < 0.01$

When we compared employees who obtained high AQ scores with employees who obtained low AQ scores, we found a very large and statistically significant difference in the average number of days absent for one year and the number of occurrences of absenteeism over that year.

Employees scoring at or above the 75th percentile in AQ consisted of those with AQ scores of 163 or higher. Employees scoring below the 25% percentile in AQ consisted of those with AQ scores of 137 or lower.

On average, members of the high scoring group missed 2.77 days of work over the course of a year, whereas members of the low scoring group missed 6.88 days of work.

On average, employees scoring among the bottom 25% in AQ had two and a half times as many days absent as did employees scoring among the top 25% in AQ, a highly significant difference (F = 17.12, p = 0.000).

Furthermore, employees in the low-scoring group had an average of 1.89 absentee occurrences compared with an average of 0.85 occurrences for the high-scoring group. These differences in absentee occurrences between high-scoring and low-scoring employees were highly significant (F = 49.16, p = 0.000). See Table 7a.

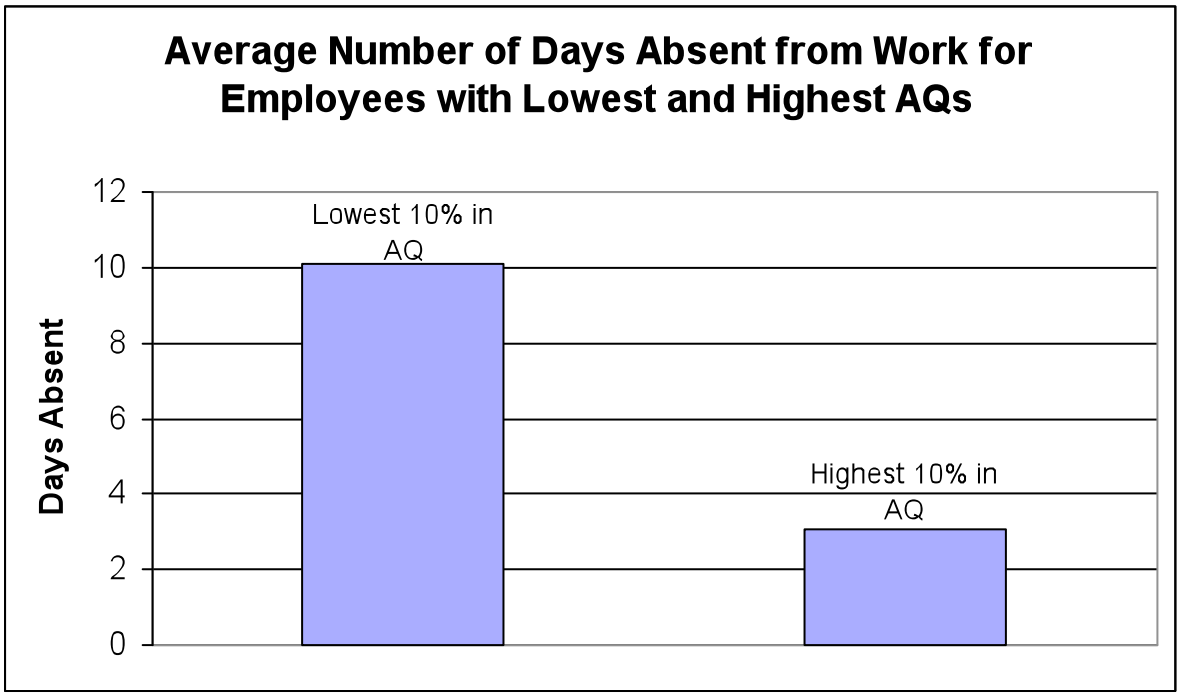
Table 7a Absentee Statistics for Subjects Scoring Above the 75th and Below the 25th Percentiles on AQ Score			
		Days Absent	Occurrences
Below 25th	Mean	6.88	1.89
	N	293	293
	Std. Dev.	15.48	2.09
Above 75th	Mean	2.77	.85
	N	285	285
	Std. Dev.	6.38	1.38

The same computations were made for employees scoring above the 90th percentile on AQ compared with those scoring below the 10th percentile. The 90th percentile corresponded to an AQ score of 175, and the 10th percentile corresponded to an AQ score of 126. Comparing these two groups, we found that on average, employees who scored in the lowest group in AQ had an average of 10.12 days absent, and employees scoring in the top 10% had an average of only 3.04 days absent, a highly significant difference ($F = 10.12$, $p = 0.002$). The average number of absentee occurrences was 2.07 for the low-scoring group and only 0.88 for the high-scoring group, also a highly significant difference ($F = 21.49$, $p = 0.000$).

Put another way, the low scoring group (AQ) had three and a third times as many days absent as the top group, and almost two and a half times as many absentee occurrences as subjects scoring in the top AQ group.

See Table 7b and accompanying graph.

Table 7b Absentee Statistics for Subjects Scoring Above the 90th and Below the 10th Percentiles on AQ Score			
		Days Absent	Occurrences
Below 10th	Mean	10.12	2.07
	N	121	121
	Std. Dev.	20.30	2.26
Above 90th	Mean	3.04	0.88
	N	119	119
	Std. Dev.	7.79	1.45



Conclusions

Results of this study indicated the following:

- The AQ score shows excellent validity as a correlate of many health, life, and work factors.
- An employee's perception of his or her health, fitness, quality of life, job satisfaction, and other health-related, happiness-related, and job-performance aspects of life are all significantly correlated to AQ.
- The AQ score demonstrates excellent validity as a reflection of employee work attendance. It does that in two ways:
    - By correlating significantly with days absent from work and with the number of absentee occurrences; the higher the AQ, the fewer absences.
    - By demonstrating that employees who score very high in AQ (among the top 10% or top 25%) have far fewer days absent, and absentee occurrences, than employees who score very low in AQ (among the bottom 10% or bottom 25%). Differences are statistically significant and quite large.
- AQ Profile subscores demonstrate excellent discriminant validity, with scale intercorrelations ranging from 0.28 to 0.72.
- Reliabilities (alpha coefficients) of AQ and the four subscores are exceptionally high (over .90 for AQ and over .80 for subscores), indicating that scores are suitable for drawing reliable inferences about individual test-takers;
- The distribution of AQ scores is very nearly normal, with a mean of 150 and a standard deviation of 18.
- Any gender difference in the AQ score, or any of its subscores, is extremely small.
- Age is slightly correlated with AQ and its subscores, suggesting that life experiences tend to improve a person's ability to respond to adversity.

Summary: A Profile of the High AQ Subject

In this study, high AQ subjects can be characterized as perceiving themselves as physically healthier, more fit, energetic, engaging in more physical exercise, eating a healthier diet, feeling less stressed, taking fewer prescription drugs, being happier, more optimistic, more successful, luckier, enduring less hardship, and enjoying their jobs more than low AQ subjects.

High AQ subjects have far fewer days absent from work as well as far fewer separate occurrences of absenteeism.

Future Directions

As more businesses make use of this latest version of the *AQP*, more information will be gathered on the validity and reliability of AQ scores. To date, there is good evidence that AQ can be a useful predictor of job performance and job attendance. More studies in the future will continue to expand on the existing research regarding which aspects of job performance are best predicted.

If your company is interested in using the *AQ Profile* for screening its applicants, developing its people, or for conducting an internal validity study, contact PEAK Learning, Inc. info@peaklearning.com (805) 595-7775 • www.peaklearning.com

References

Abramson, L.Y., Seligman, M.E.P., & Teasdale, J.D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74.

Aldwin, C.M., & Gilman, D.F. (2004). *Health, illness and optimal ageing.* London: Sage.

Campbell, D.T. (1960). Recommendations for the APA test standards regarding construct, trait, and discriminant validity. *American Psychologist, 15*, 546-553).

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Erlbaum.

Compton, W.C. (2005). An introduction to positive psychology. Belmont, CA: Thomson-Wadsworth.

Grandy, J. (2000). *Psychometric assessment of the Adversity Response Profile*. (Unpublished manual).

Hiroto, D.S., & Seligman, M.E.P. (1975). Generality of learned helplessness in man. *Journal of Personality and Social Psychology, 37*, 1-11.

Irwin, M. & Vedhara, K. (2005). *Human Psychoneuroimmunology*. Oxford University Press.

Kobasa, S.C. (1979). Stressful life events, personality and health: An inquiry into hardiness. *Journal of Personality and Social Psychology, 42*, 168-177.

Ong, A.D., Bergeman, C.S., Bisconti, T.L., & Wallace, K.A. (2006). Psychological resilience, positive emotions, and successful adaption to stress in later life. *Journal of Personality and Social Psychology, 91(4)*, 730-749.

Peterson, C., Maier, S.F., & Seligman, M.E.P. (1993). *Learned helplessness: A theory for the age of personal control*. New York: Oxford University Press.

Peterson, C. & Seligman, M.E.P. (2004). Character strengths and virtues: A handbook and classification. Oxford: Oxford University Press.

Reivich, K. & Shatte, A. (2002). The resilience factor: 7 keys to finding your inner strength and overcoming life's hurdles. New York: Broadway Books.

Rotter, J. (1966). Generalized expectancies for internal versus external locus of control of reinforcement. *Psychological Monographs, 81*(1, Whole No. 609).

Schneider, L.M., & Briel, J.B. (1990). *Validity of the GRE: 1988-89 Summary Report*. Princeton, NJ: Educational Testing Service.

Seligman, M.E.P. (1991). *Learned optimism*. New York: Knopf.

Stoltz, P.G. (1997). *Adversity quotient: Turning obstacles into opportunities*. New York: Wiley.

Stoltz, P.G. (2000). *Adversity quotient @ work: Make everyday challenges the key to your success—Putting the principles of AQ into action*. New York: Morrow.

Stoltz, P.G. & Weihenmayer, E. (2008). *The adversity advantage: Turning everyday struggles into everyday greatness.* New York: Simon & Schuster.

Tugade, M.M., Frederickson, B.L., & Barrett, L.F. (2004). Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of Personality, 72(6)*, 1161-1190.

Wortman, C.B., & Brehm, J.W. (1975). Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8). New York: Academic Press.

Terminology Used in This Report	
***AQ Profile* Terms**	
AQ Profile	The instrument used to measure an individual's style of responding to adverse situations
Adversity Quotient® (AQ)®	The total score obtained on the AQ Profile.
Control score	A measure of the degree of control a person perceives that he or she has over adverse events; a scale on the AQ Profile and a component of the Adversity Quotient
Ownership score	A measure of the extent to which a person owns, or takes responsibility for, the outcomes of adversity or the extent to which a person holds himself or herself accountable for improving the situation
Reach score	A measure of the degree to which a person perceives good or bad events reaching into other areas of life
Endurance score	A measure of the perception of time over which good or bad events and their consequences will last or endure
Statistical terms	
Mean	The common average obtained by adding up everyone's score and dividing by the number of people
Norms	A distribution of scores obtained by a defined sample of people
Psychometric terms	
Coefficient alpha	A measure of internal-consistency reliability ranging from 0 to 1
Convergent validity	Demonstration that a test measures what it is purported to measure
Discriminant validity	Demonstration that a test does not measure traits or knowledge other than what it is purported to measure
Reliability	The consistency with which people give the same answers to questions or to similar questions
Validity	The degree to which a test measures what it purports to measure and does not measure other traits, knowledge, or skills