

MANUEL TECHNIQUE & GUIDE UTILISATEUR GENERAL ANALYTICAL APTITUDE TEST (GAAT)

Ce manuel a été réalisé par les psychologues du département de Recherche et Innovation de Central Test.

*Le test GENERAL ANALYTICAL APTITUDE TEST (**GAAT**) et le logo Central Test sont des marques déposées.*

Tout droit de reproduction, modification ou réutilisation doit être accordé par écrit par l'éditeur Central Test.

*Copyright © 2020 par Central Test - Empowering Talent.
Tous droits réservés.*

Sommaire

Introduction	3
1. Description générale et domaines d'application	4
1.1. Descriptions des facteurs	4
1.2. Principaux domaines d'application	5
1.3. Administration et structure du questionnaire	6
1.4. Questions techniques et de sécurité	6
2. Propriétés psychométriques	8
2.1. Description de l'échantillon	8
2.2. Normes	9
2.3. Fidélité	13
2.4. Validité	13
2.5. GAAT et égalité des chances	15
3. Informations et considérations supplémentaires	19
Références	20

Introduction

Comment les individus pensent-ils ? Quels mécanismes utilisent-ils pour résoudre les problèmes et ces mécanismes peuvent-ils être appris ? Bien que de nombreux aspects de la cognition humaine demeurent un mystère, les psychologues ont commencé à apporter plus de précisions sur les stratégies utilisées pour penser de manière organisée, analyser et résoudre des problèmes. Ce style de pensée est généralement appelé « pensée critique ».

La pensée critique a été étudiée dans les années 1910, lorsque John Dewey a publié son livre phare *How We Think* (1910/1933). Elle fait partie de nombreux modèles de compétences jugés importants pour la réussite du personnel d'une entreprise (Trilling & Fadel, 2009). Le raisonnement, la logique et le jugement sont largement admis comme étant des aptitudes cognitives utiles sur le lieu de travail et des éléments importants de la pensée critique. Le développement d'une évaluation de la pensée critique était alors nécessaire pour aider les organisations à évaluer et recruter le bon candidat, c'est pourquoi le test d'aptitude analytique générale - General Analytical Aptitude Test (**GAAT**) - a été créé.

1. Description générale et domaines d'application

La pensée critique est définie comme la capacité de juger une situation complexe ou un argument de manière réfléchie et résolue. Ce jugement est généralement basé sur des données clairement identifiées et permet à un individu de parvenir à une conclusion logique et de décider quoi croire ou quoi faire. Dans n'importe quel contexte, l'utilisation de la pensée critique implique de poser des questions telles que : Quel est le cœur du problème? Quelle approche devrais-je utiliser pour le résoudre? Est-ce que ces données me permettront bien d'arriver à une conclusion? Cet argument n'est-il pas en opposition avec cet autre argument ? Si cela est vrai, qu'est-ce que cela implique? Etc.

Par conséquent, le test **GAAT** est conçu pour évaluer les capacités de pensée critique d'un individu au moyen de quatre facteurs cruciaux : l'évaluation des arguments, l'analyse critique, le raisonnement déductif et le raisonnement inductif.

GAAT se distingue des autres outils similaires sur le marché car il possède les avantages suivants :

- Il résume la performance globale de la personne au moyen d'un quotient de pensée critique ;
- Il présente une note globale sur une échelle de zéro à vingt, comprenant le nombre de réponses correctes ;
- Le rapport présente une interprétation personnalisée du score global de la personne ;
- On trouve une description détaillée de la pensée critique, de son histoire et de sa mesure dans le **GAAT** ;
- Il propose une représentation graphique des scores pour chacun des quatre facteurs (Evaluation des arguments, Analyse critique, Raisonnement déductif et Raisonnement inductif) en référence à la population.

1.1. Descriptions des facteurs

Le test **GAAT** est conçu de manière à mesurer le processus d'analyse, de synthèse, de conceptualisation, d'application et/ou d'évaluation de l'information provenant de diverses sources au moyen de quatre éléments importants, à savoir *l'évaluation des arguments, l'analyse critique, le raisonnement déductif et le raisonnement inductif*.

Evaluation des arguments - L'évaluation des arguments correspond à la faculté de comprendre, formuler et déconstruire des arguments, faire la distinction entre des informations importantes et superficielles.

Lorsque nous cherchons à convaincre les autres, nous utilisons différents arguments afin d'appuyer nos points de vue. Un argument efficace devrait amener les autres à nos conclusions sur la base d'un accord avec les informations que nous apportons. Connaître l'ordre dans lequel présenter habilement des arguments peut rendre notre discours plus convaincant. Lors d'un débat, il est essentiel d'être capable de répondre à vos opposants en déconstruisant efficacement leurs arguments, en identifiant et en mettant en évidence la faiblesse de leurs affirmations.

Ces compétences ne sont pas seulement importantes dans un grand nombre de métiers, comme la vente ou le droit, mais également utiles au quotidien, par exemple, lorsque l'on conduit une négociation ou que l'on collabore ou débat sur divers opinions et sujets avec autrui.

Analyse critique - L'analyse critique correspond à la faculté de comprendre et d'interpréter objectivement et systématiquement les informations afin d'en tirer des conclusions raisonnées.

Il est important d'utiliser cette compétence pour analyser les connexions logiques entre des idées et de les évaluer efficacement afin de faire un jugement sur leur valeur. Ceci nous permet de prendre des décisions quant à nos actions sur la base d'options diverses et variées.

L'analyse critique est une compétence que nous utilisons régulièrement, par exemple, lors de conversations, lorsque nous lisons et jugeons de la fiabilité d'une information, ou encore lorsque nous résolvons des problèmes ou que nous réfléchissons à nos actions au quotidien.

Raisonnement déductif - Le raisonnement déductif correspond à la faculté de tirer certaines conclusions logiques à partir de lois parmi de nombreuses prémisses. Ce type de raisonnement est communément appliqué dans l'utilisation de syllogismes, ou de déclarations de type "si-alors". Si $A=B$ et $B=C$, alors un raisonnement déductif nous permet de conclure logiquement que $A=C$.

Si les arguments déductifs sont créés avec précision, le raisonnement déductif nous permet de tirer des conclusions qui doivent être toujours vraies. C'est pour cela qu'un argument doit être à la fois valide d'un point de vue déductif (les conclusions sont vraies aussi longtemps que les prémisses sont vraies) et valable (le contenu des prémisses est vrai).

Nous utilisons ce type de raisonnement dans de nombreux champs professionnels, comme pour trouver des applications concrètes aux lois de la physique et de la biologie dans la production de médicament, de stratégies marketing ou de pratiques dans les ressources humaines.

Raisonnement inductif - Le raisonnement inductif correspond à la capacité de faire des inférences et des généralisations sur la base des prémisses spécifiques d'une information. Par exemple, étant donné que tous les cygnes que j'ai vus sont blancs, je généralise et considère que tous les cygnes sont blancs.

Nous utilisons également ce type de raisonnement lorsque nous prenons des décisions sur la base d'informations issues de notre expérience. En déterminant les liens entre des données à l'aide des connaissances issues de nos expériences passées, nous arrivons à des conclusions générales à propos de l'information en utilisant un raisonnement inductif.

Le raisonnement inductif est fréquemment utilisé dans notre vie professionnelle et quotidienne, généralement dans la plupart des situations qui impliquent des prévisions ou des pronostics, utilisant nos connaissances pour prendre des décisions éclairées. Ce type de raisonnement est capital en science et en recherche, où les scientifiques recueillent des données, expérimentent et font des hypothèses pour des recherches plus poussées.

1.2. Principaux domaines d'application

Plusieurs études, notamment celle de Schmidt et Hunter (1996), montrent que la capacité mentale générale (intelligence) est l'un des principaux prédicteurs du rendement au travail. En effet, la recherche affirme que cette capacité à comprendre des choses plus ou moins complexes et ce, de manière rapide, ainsi que la capacité à continuellement adapter et améliorer son expertise sont de bons prédicteurs. Par conséquent, dans le domaine professionnel, il est impératif de mesurer la pensée critique car c'est une qualité indispensable pour prendre rapidement les bonnes décisions.

Le **GAAT** évalue les capacités de l'individu en le plaçant dans des situations proches de la réalité. La compréhension précise et l'analyse des questions du test reflètent la manière dont il perçoit les informations dans des situations réelles. Ainsi, c'est une évaluation directe de la capacité à comprendre et à prendre des décisions.

Par conséquent, le **GAAT** fournit une évaluation objective et fondée sur des normes du niveau d'aptitude de la personne à la pensée critique, tel qu'elle est utilisée dans un contexte professionnel. Ce test est recommandé pour le recrutement d'employés dont les fonctions supposent l'utilisation de la pensée critique et des prises de décision. **GAAT** s'applique également au recrutement et à la pré-formation de personnes dans des postes administratifs, de manager et/ou de direction, ainsi que pour les étudiants participant à des examens d'entrée ou à des organisations, en prélude à des programmes de développement personnel.

Le **GAAT** est spécialement réservé aux professionnels des ressources humaines des organisations, des écoles et des collèges, ainsi qu'aux formateurs.

1.3. Administration et structure du questionnaire

Administration à distance du GAAT

Le **GAAT** peut être administré dans des contextes non supervisés. Cependant, l'administrateur du test doit prendre les mesures appropriées pour minimiser les risques d'obtenir une aide externe de la part du candidat.

Format du GAAT :

- Nombre de questions : 30
- Format du test : questionnaire à choix multiples
- Durée : 40 minutes (chronométré)
- Administration : En ligne

1.4. Questions techniques et de sécurité

Les points suivants servent de directives pour l'administration sécurisée du **GAAT**. Celles-ci sont utiles pour veiller à ce que le **GAAT** soit administré de manière objective et équitable. Les spécifications de sécurité sont également importantes pour la protection des informations du candidat, à la fois pendant l'administration et après l'évaluation.

Respect de la charte éthique

En tant qu'éditeur d'outils d'évaluation en ligne, *Central Test* respecte les conventions juridiques du droit du travail français et les dispositions réglementaires issues de la loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés : la déclaration n ° 1834067. De plus, depuis le 25 mai 2018, *Central Test* est conforme aux réglementations de la RGPD (règlement général sur la protection des données) édictées par le Parlement et le Conseil européen.

Les résultats du test et les données personnelles ne peuvent pas être partagés avec des tiers sans le consentement explicite du candidat. Les candidats ont le droit d'accéder à leurs résultats et l'évaluateur doit leur accorder l'accès sur demande. De plus, si le candidat le demande, toutes les données doivent être définitivement supprimées. Pour ce faire, le candidat doit contacter directement l'organisation ou *Central Test*.

Confidentialité et sécurité

Central Test garantit la sécurité et la confidentialité des données générées lors de l'administration des tests en ligne. En effet, les codes d'accès sont cryptés dans la base de données et ne sont accessibles à personne. Les mots de passe sont générés par ordinateur et ne sont pas visibles pour les entreprises. Un module de réécriture d'URL (générateur de lien sécurisé) a été installé et dispose d'un paramètre transparent pour l'utilisateur.

La navigation et les données échangées via l'application de *Central Test* peuvent être sécurisées et représentées par une icône sous la forme d'un cadenas dans le navigateur Internet. L'application de *Central Test* dispose d'une signature numérique << GeoTrust SSL TRUE BusinessID >>, émise et certifiée par GeoTrust®. Ce certificat protège toutes les évaluations de *Central Test* des attaques de phishing et constitue la preuve de notre engagement à fournir une plate-forme sécurisée.

Les invitations à passer des tests se font via des liens sécurisés d'utilisation limitée (en nombre et en durée). Les tests sont chronométrés et les adresses IP sont suivies.

2. Propriétés psychométriques

Chez *Central Test*, les tests sont continuellement soumis à révision afin d'assurer leur applicabilité et leur efficacité. C'est également le cas du **GAAT**. Le questionnaire a été prétesté à plusieurs reprises pour s'assurer qu'il répondait aux critères de fidélité et de validité établis par des institutions psychologiques reconnues sur le plan international. Après deux prétests, le dernier prétest comprend 2261 professionnels francophones (hommes et femmes) d'environ 40 pays. Les échantillons sont composés de francophones de différents continents afin d'établir le questionnaire comme mesure internationale de la pensée critique, ainsi que pour s'assurer que les items sont précis, pertinents, clairs et concis et qu'ils reflètent le construit ciblé.

2.1. Description de l'échantillon

Tableau 1. Échantillon par genre.

Genre	Pourcentage (%)	N
Femme	30,48	694
Homme	68,82	1567

Tableau 2. Échantillon par groupe d'âge.

Groupe d'âge	Pourcentage (%)	N
Moins de 18 ans	2,33	53
18-25 ans	19,98	455
26-33 ans	37,68	858
34-41 ans	22,97	523
42-49 ans	11,20	255

50-56 ans	2,72	62
Plus de 57 ans	1,05	24

Tableau 3. Échantillon par niveau d'études.

Niveau d'étude	Pourcentage (%)	N
En cours d'études	2,59	59
Niveau 1 : Diplôme de 2e ou 3e cycle universitaire / enseignement supérieur	59,42	1,353
Niveau 2 : Diplôme de 1er cycle universitaire / enseignement supérieur	13,79	314
Niveau 3 : Diplôme d'études professionnelles ou techniques, BTS, DUT	9,84	224
Niveau 4 : BAC / Diplôme d'études secondaires	6,10	139
Niveau 5 : Qualification professionnelle, CAP, BEP	4,35	99
NC	3,91	89

2.2. Normes

Les normes nous aident à comparer le score d'un individu à un test avec les scores d'autres personnes ayant effectué le même test, généralement dans des conditions similaires. Ces personnes sont représentatives de la population pour laquelle une évaluation est prévue et sont donc appelées « groupe de référence ». Plus la correspondance entre les caractéristiques de l'échantillon et de la population cible sont proches, plus la distribution sera précise en tant que référence de notation. Grâce aux normes, les développeurs de tests et les psychologues sont capables d'établir des comparaisons entre les individus et de prendre des décisions plus objectives. Les groupes de normes nous permettent d'attribuer un sens aux résultats des tests et de quantifier la performance d'un individu, le niveau d'un attribut ou d'un trait par rapport à ses pairs.

Les tableaux 3 à 8 détaillent les scores moyens (moy), l'écart-type (σ), les scores minimum (Min) et maximum (Max), l'intervalle et l'estimation standard (SE) pour l'échantillon francophone sur les facteurs de **GAAT** ainsi que pour le quotient de pensée critique. Ces tableaux sont divisés en catégories selon l'âge, le genre et le niveau d'études.

Tableau 3. Quotient de Pensée critique pour l'ensemble de l'échantillon.

Quotient	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Pensée critique	96,40	13,41	93,83	74,13	141,7	67,57	0,28

Tableau 4. Scores par facteurs pour tout l'échantillon.

Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Analyse critique	5,02	1,72	4,38	3,05	10,0	6,95	0,04
Raisonnement déductif	5,10	1,81	5,12	2,75	10,0	7,25	0,04
Évaluation des arguments	5,33	1,93	4,66	2,15	10,0	7,85	0,04
Raisonnement inductif	5,14	1,76	5,24	3,10	10,0	6,90	0,04

Tableau 5. Quotient de pensée critique en fonction de l'âge.

Quotient	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Moins de 18 ans	104,24	17,12	105,50	74,13	141,7	67,57	2,35
18-25 ans	97,02	12,36	94,30	74,13	136,85	62,72	0,58
26-33 ans	94,22	12,15	92,01	74,13	141,7	67,57	0,41
34-41 ans	96,51	13,79	93,76	74,13	141,7	67,57	0,60
42-49 ans	98,56	15,18	95,91	74,13	141,7	67,57	0,95
50-56 ans	101,86	14,48	99,91	76,95	138,81	61,86	1,84
57+	96,76	16,90	93,49	74,13	141,7	67,57	3,45

Tableau 6. Scores par facteur en fonction du groupe d'âge.

Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Moins de 18 ans							
Analyse critique	6,08	2,22	5,81	3,05	10,0	6,95	0,30
Raisonnement déductif	6,00	2,09	6,36	2,75	10,0	7,25	0,29
Évaluation des arguments	5,79	2,40	5,87	2,15	10,0	7,85	0,33
Raisonnement inductif	6,08	2,13	6,35	3,10	10,0	6,90	0,29
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
18 – 25 ans							
Analyse critique	5,01	1,65	4,38	3,05	9,79	6,74	0,08
Raisonnement déductif	5,33	1,76	5,12	2,75	10,00	7,25	0,08
Évaluation des arguments	5,35	1,89	4,66	2,15	10,00	7,85	0,09
Raisonnement inductif	5,17	1,68	5,24	3,10	10,00	6,90	0,08
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
26 – 33 ans							
Analyse critique	4,89	1,63	4,38	3,05	10,0	6,95	0,06
Raisonnement déductif	4,85	1,68	5,12	2,75	10,0	7,25	0,06
Évaluation des arguments	5,00	1,80	4,66	2,15	10,0	7,85	0,06
Raisonnement inductif	4,91	1,64	4,21	3,10	10,0	6,90	0,06
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
34 – 41 ans							
Analyse critique	4,97	1,75	4,38	3,05	10,0	6,95	0,08
Raisonnement déductif	5,08	1,86	5,12	2,75	10,0	7,25	0,08
Évaluation des arguments	5,44	1,91	4,66	2,15	10,0	7,85	0,08
Raisonnement inductif	5,15	1,77	5,24	3,10	10,0	6,90	0,08

Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
42 – 49 ans							
Analyse critique	5,27	1,85	4,38	3,05	10,0	6,95	0,12
Raisonnement déductif	5,22	1,99	5,12	2,75	10,0	7,25	0,12
Évaluation des arguments	5,61	2,14	5,87	2,15	10,0	7,85	0,13
Raisonnement inductif	5,42	1,88	5,24	3,10	10,0	6,90	0,12
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
50 – 56 ans							
Analyse critique	5,25	1,80	5,81	3,05	9,79	6,74	0,23
Raisonnement déductif	5,38	1,75	5,12	2,75	9,87	7,12	0,22
Évaluation des arguments	6,38	1,84	5,87	3,36	10,00	6,64	0,23
Raisonnement inductif	5,92	1,98	6,35	3,10	10,00	6,90	0,25
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Plus de 57 ans							
Analyse critique	5,15	1,90	4,38	3,05	10,0	6,95	0,39
Raisonnement déductif	5,10	2,18	3,98	2,75	10,0	7,25	0,45
Évaluation des arguments	5,63	2,40	5,27	2,15	10,0	7,85	0,49
Raisonnement inductif	4,87	2,07	4,21	3,10	10,0	6,90	0,42

Tableau 7. Quotient de pensée critique en fonction du genre.

Quotient	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Femmes	98,74	14,27	96,44	74,13	141,7	67,57	0,54
Hommes	95,23	12,72	92,45	74,13	141,7	67,57	0,32

Tableau 8. Scores par facteur en fonction du genre.

Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Femmes							
Analyse critique	5,21	1,79	4,38	3,05	10,0	6,95	0,07
Raisonnement déductif	5,43	1,93	5,12	2,75	10,0	7,25	0,07
Évaluation des arguments	5,53	2,02	5,87	2,15	10,0	7,85	0,08
Raisonnement inductif	5,42	1,92	5,24	3,10	10,0	6,90	0,07
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Hommes							
Analyse critique	4,92	1,67	4,38	3,05	10,0	6,95	0,04
Raisonnement déductif	4,94	1,72	5,12	2,75	10,0	7,25	0,04
Évaluation des arguments	5,23	1,88	4,66	2,15	10,0	7,85	0,05
Raisonnement inductif	5,00	1,65	4,21	3,10	10,0	6,90	0,04

Tableau 9. Quotient de pensée critique en fonction du niveau d'études

Quotient	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Niveau 1 : Diplôme de 2e ou 3e cycle universitaire / enseignement supérieur	96,48	13,38	94,04	74,13	141,7	67,57	0,36
Niveau 2 : Diplôme de 1er cycle universitaire / enseignement supérieur	95,11	12,52	91,91	74,13	136,92	62,79	0,71
Niveau 3 : Diplôme d'études professionnelles ou techniques, BTS, DUT	96,51	12,82	96,18	74,13	134,26	60,13	0,86
Niveau 4 : BAC / Diplôme d'études secondaires	90,99	11,18	88,27	74,13	130,99	56,86	0,95
Niveau 5 : Qualification professionnelle, CAP, BEP	96,26	14,41	91,21	74,13	137,29	63,16	1,45
En cours d'études	101,60	13,91	99,76	76,95	133,37	56,42	1,81

Tableau 10. Scores par facteur en fonction du niveau d'études

Level 1: Master's Degree, PhD, Engineering or Business School, MBA							
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Analyse critique	5,03	1,71	4,38	3,05	10,0	6,95	0,05
Raisonnement déductif	5,11	1,79	5,12	2,75	10,0	7,25	0,05
Évaluation des arguments	5,36	1,93	4,66	2,15	10,0	7,85	0,05
Raisonnement inductif	5,12	1,75	5,24	3,10	10,0	6,90	0,05

Level 2: Bachelor's Degree, First-year Masters							
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Analyse critique	4.97	1.69	4.38	3.05	10.00	6.95	0.10
Raisonnement déductif	4.94	1.76	5.12	2.75	10.00	7.25	0.10
Évaluation des arguments	5.12	1.89	4.66	2.15	9.59	7.44	0.11
Raisonnement inductif	5.01	1.65	4.21	3.10	10.00	6.90	0.09

Level 3: 2 years of vocational or university studies							
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Analyse critique	5.00	1.72	4.38	3.05	10.00	6.95	0.11
Raisonnement déductif	5.06	1.76	5.12	2.75	9.87	7.12	0.12
Évaluation des arguments	5.39	1.82	5.87	2.15	10.00	7.85	0.12
Raisonnement inductif	5.19	1.74	5.24	3.10	10.00	6.90	0.12

Level 4: high school diploma							
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Analyse critique	4.58	1.41	4.38	3.05	8.46	5.41	0.12
Raisonnement déductif	4.42	1.62	3.98	2.75	9.87	7.12	0.14
Évaluation des arguments	4.63	1.79	4.66	2.15	9.59	7.44	0.15
Raisonnement inductif	4.65	1.48	4.21	3.10	10.00	6.90	0.13

Level 5: trades training, no high school diploma							
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Analyse critique	5.12	1.96	4.38	3.05	10.00	6.95	0.20
Raisonnement déductif	5.11	1.90	5.12	2.75	9.87	7.12	0.19
Évaluation des arguments	5.19	2.10	4.66	2.15	9.59	7.44	0.21
Raisonnement inductif	5.11	1.76	5.24	3.10	10.00	6.90	0.18

En cours d'études							
Facteur	Moyenne	SD	Médiane	Min	Max	Etendue	SE
Analyse critique	5.28	1.92	4.38	3.05	9.79	6.74	0.25
Raisonnement déductif	5.88	1.89	5.12	2.75	10.00	7.25	0.25
Évaluation des arguments	5.93	1.79	5.87	2.15	10.00	7.85	0.23
Raisonnement inductif	5.74	1.95	5.24	3.10	10.00	6.90	0.25

2.3. Fidélité

La fidélité indique dans quelle mesure une évaluation ne comprend pas d'erreur et le degré auquel les items d'une échelle mesurent le même construit. Par exemple, tous les items qui évaluent l'égalité sont cohérents et reflètent la définition de celle-ci. Une première préoccupation, lorsque l'on élabore et utilise les évaluations, est de savoir à quel point les erreurs de la mesure aléatoire influencent la performance du test. Si un test n'est pas fiable, il est impossible de le valider car l'exactitude des résultats (validité) dépend de la cohérence (fidélité) avec laquelle ces mêmes résultats sont produits. Une haute fidélité permet d'éviter les biais pouvant être causés par des erreurs de jugement humain lors du processus de recrutement et de sélection.

Les valeurs de fidélité interne varient de 0 à 1, les valeurs les plus élevées étant plus souhaitables. En règle générale, une fidélité de 0,70 ou plus est recommandée (Nunnally, 1978), avec des valeurs supérieures à 0,9 (excellent), 0,7 à 0,9 (bon), 0,6 à 0,7 (acceptable); 0,5 à 0,6 (faible) et des valeurs inférieures à 0,5 (inacceptable).

Le **GAAT** a une fidélité globale de 0,82, ce qui indique que la balance a une bonne cohérence interne et une grande fiabilité.

2.4. Validité

La validité se réfère à la mesure dans laquelle une évaluation mesure ce qu'elle prétend mesurer, que cela est bien fondé et correspond avec précision au comportement réel. En termes de validité psychométrique, cela se réfère au degré auquel les faits et la théorie scientifique étayent l'interprétation des scores aux tests. Alors que la fidélité d'une évaluation est relative à la mesure dans laquelle les résultats sont cohérents dans le temps, la validité nous indique dans quelle mesure les résultats d'une évaluation sont précis et nous permettent de faire des prédictions fiables sur un résultat. Par conséquent, un test peut être fidèle mais non valide (un participant peut obtenir la mauvaise réponse de façon systématique) mais il ne peut être valide sans être fidèle, car la confiance dans la précision des résultats dépend de la cohérence avec laquelle on obtient les mêmes résultats. La validité établie protège également contre les décisions personnelles fautives, coûteuses, non éthiques et juridiquement risquées.

Les facteurs du **GAAT** évaluent la pensée critique, mais ils mesurent tous différents facteurs de cette compétence. En d'autres termes, si une personne obtient un score élevé sur un facteur, cela ne signifie pas qu'elle obtiendra un score élevé sur tous les facteurs. Le coefficient de corrélation est une façon numérique de quantifier la relation entre deux variables et se situe entre -1 et 1. Plus la valeur est proche de 1, plus la relation est forte. Le coefficient de corrélation n'implique pas de causalité; cela signifie simplement l'association entre deux variables. Traditionnellement, les corrélations de facteurs sont mesurées à l'aide de la corrélation de Pearson (1895).

Tableau 9. Corrélations Inter-facteur.

Facteur	Quotient	Analyse critique	Raisonnement déductif	Évaluation des arguments	Raisonnement inductif
Quotient	1				
Analyse critique	0,72	1			

Raisonnement déductif	0,82	0.48	1		
Évaluation des arguments	0,80	0.39	0,55	1	
Raisonnement inductif	0,83	0.47	0,59	0,57	1

Les résultats présentés dans le tableau 9 montrent qu'il existe une relation linéaire positive entre les quatre facteurs - *Évaluation des arguments*, *Analyse critique*, *Raisonnement déductif* et *Raisonnement inductif*.

Une corrélation positive est à prévoir dans une évaluation telle que le **GAAT** dans la mesure où les quatre facteurs contribuent à la mesure d'un facteur de quotient de second ordre ou, en d'autres termes, d'une capacité de pensée critique globale.

2.5. GAAT et égalité des chances

Les organisations sont de plus en plus sous pression pour s'assurer de pouvoir montrer que les outils qu'elles utilisent dans leurs processus de recrutement, sélection, développement, promotion et évaluation des performances des employés sont à la fois justes, objectifs et précis. Cette pression découle du fait que les candidats sont de plus en plus informés. De plus, la main-d'œuvre est de plus en plus exigeante et la législation locale et internationale ainsi que les normes concernant les tests sont de plus en plus strictes en matière d'emploi. L'égalité des chances représente donc une considération essentielle pour toute évaluation du travail moderne, et une évaluation n'est considérée comme réellement valable que lorsqu'elle est appropriée pour être utilisée avec différents groupes. **GAAT** est une mesure juste et précise de la pensée critique, et les praticiens des RH et les psychologues peuvent avoir confiance en ses applications diverses.

Équivalence de mesure

L'équivalence de mesure désigne la cohérence de la mesure entre différents groupes aux caractères spécifiques, tels que la nationalité, l'origine ethnique et le genre. Elle vise à garantir que les personnes de différents groupes ayant le même statut sur un attribut aient le même ensemble d'observations lors d'une évaluation (Drasgow, 1987; McLeary & Cruise, 2012). Elle aide à déterminer dans quelle mesure les répondants interprètent les items de la même manière, et à savoir si les différences identifiées entre les groupes reflètent les différences réelles dans le caractère mesuré (Drasgow, 1987). L'analyse différentielle du fonctionnement des items est une des approches permettant d'évaluer l'équivalence de mesure.

Analyse différentielle du fonctionnement des items (Dif)

Un item fonctionne différemment, ou affiche Dif, si les répondants de différents groupes ont des probabilités d'endosser l'item différentes, bien qu'ils aient le même niveau sur un attribut ou un trait. Quand une probabilité différente d'endosser un item reste constante à travers différents niveaux de l'attribut ou trait, le Dif est dit "uniforme". Dans ce cas, l'item concerné donne systématiquement un avantage à un groupe à tous les niveaux sur l'attribut ou trait. Inversement, quand une probabilité différente d'endosser un item varie au niveau de l'attribut ou du trait, le Dif est qualifié de "non uniforme" (Clauser et Mazor, 1998; Hanson, 1998). Plutôt qu'un avantage constant accordé au groupe de référence à travers le continuum des capacités, un avantage pour un groupe par

rapport à un autre existe pour différents niveaux du trait. Par exemple, un item peut avantager le groupe de référence à l'extrémité supérieure du continuum, tout en le désavantageant à l'extrémité inférieure.

Analyse DIF : le genre

L'analyse DIF a été réalisée selon le genre afin de s'assurer qu'il n'y avait pas de préjugés envers les hommes ou les femmes, et donc que **GAAT** est un test équitable pour les deux genres. Pour cette analyse, l'échantillon de normes a été divisé en fonction du genre (694 femmes et 1567 hommes).

Tableau 11. Analyse DIF et taille de l'effet pour les items DIF.

Factor	Item No.	Chisq-value	P-value	Taille de l'effet
Evaluation des arguments	13	13,931	0,000	Négligeable
Evaluation des arguments	27	55,930	0,000	Négligeable
Analyse critique	9	8,4122	0,015	Négligeable
Raisonnement déductif	10	53.559	0,000	Négligeable
Raisonnement déductif	19	12.473	0,002	Négligeable
Raisonnement déductif	25	8.0282	0,018	Négligeable
Raisonnement déductif	26	14.913	0,001	Négligeable
Raisonnement déductif	29	12.694	0,002	Négligeable
Raisonnement inductif	5	9,9807	0,007	Négligeable
Raisonnement inductif	14	19,244	0,000	Négligeable
Raisonnement inductif	17	8,9432	0,011	Négligeable

L'analyse a identifié 11 des 30 items DIF, en fonction du genre. Le tableau 11 indique le coefficient DIF ainsi que la taille de l'effet R^2 . Selon la classification de Zumbo et Thomas (1997) de la taille de l'effet, tout R^2 inférieur à 0,13 est considéré comme négligeable. Sur les 7 items identifiés comme items DIF, aucun ne présente une taille d'effet supérieur ou égal à 0,13. Par conséquent, on peut considérer que le test **GAAT** ne contient aucun élément fonctionnant différemment selon le genre.

T-tests et ANOVAS

L'analyse du DIF montre que le test **GAAT** a une validité élevée et que bien que 3 éléments aient été signalés par l'analyse DIF pour le genre, leur impact sur l'évaluation globale est négligeable, ce qui signifie que le test ne privilégie pas un genre par rapport à l'autre. Par conséquent, toute différence significative peut être attribuée à

de réelles différences entre les groupes, plutôt qu'à des biais de l'évaluation. Pour explorer davantage les différences trouvées entre les groupes, un test t d'échantillons indépendants a été réalisé entre les groupes. Les tests t et ANOVAS ont été réalisés pour le genre, le groupe d'âge et le niveau de l'étude (voir tableaux 15 à 16).

Analyse selon le genre

Afin d'étudier les différences de scores en fonction du genre, un test t d'échantillons indépendants a été réalisé pour comparer tous les facteurs du **GAAT** entre les femmes (n = 694) et les hommes (n = 1567) (voir tableau 12).

Tableau 12. T-tests par facteur, selon le genre.

Facteur	Moyenne Femmes	Moyenne Hommes	T,	df	P
Évaluation des arguments	5,46	5,20	2,89	2142	0,00
Analyse critique	5,18	4,90	3,55	2142	0,00
Raisonnement déductif	5,34	4,91	5,14	2142	0,00
Raisonnement inductif	5,33	4,97	4,45	2142	0,00

Les tests t ont révélé des différences significatives entre les hommes et les femmes pour tous les facteurs, les femmes obtenant une note légèrement supérieure à celle des hommes.

Analyse selon le groupe d'âge

Afin de vérifier les différences de scores entre les groupes d'âge, un test t avec des échantillons indépendants a été réalisé pour comparer tous les facteurs du **GAAT** entre quatre groupes d'âge: Moins de 18 ans (n=53), 18- 25 (n = 455), 26-33 (n = 858), 34- 41 (n = 523), 42-49 (n = 255), 50-56 (n = 62) et plus de 57 ans (n = 24) (voir tableau 13).

Tableau 13. ANOVA selon le groupe d'âge.

Facteur	>18	18-25	26-33	34-41	42-49	50-56	57<	F	df	p-value
Analyse critique	6,08	5,01	4,89	4,97	5,27	5,25	5,15	2,576	5	0,02
Raisonnement déductif	6,00	5,33	4,85	5,08	5,22	5,38	5,10	6,666	5	0,00
Évaluation des arguments	5,79	5,35	5,00	5,44	5,61	6,38	5,63	14,27	5	0,00
Raisonnement inductif	6,08	5,17	4,91	5,15	5,42	5,92	4,87	8,451	5	0,00

L'ANOVA a révélé des différences significatives pour tous les facteurs. Une analyse post-hoc a été réalisée à l'aide de tests t pairwise (correction de Bonferroni) afin de voir où ces différences se produisaient.

Pour l'*Analyse critique*, les personnes de 42 à 49 ans ($M = 5,27$) ont obtenu un score supérieur à celui des personnes de 26 à 33 ans ($M = 4,89$, $p = 0,0019$) et de 34 à 41 ans ($M = 4,97$, $p = 0,0209$).

En *raisonnement déductif*, les individus âgés de 18 à 25 ans ($M = 5,33$) ont obtenu un score plus élevé que les individus âgés de 34 à 41 ans ($M = 5,08$, $p = 0,03226$). Les individus de 26 à 33 ans ($M = 4,85$) ont obtenu un score inférieur à ceux de 34 à 41 ans ($M = 5,08$, $p = 0,01976$), 42 à 49 ans ($M = 5,22$, $p = 0,0037$) et 50 à 56 ans. ($M = 5,38$, $p = 0,0006$).

En *évaluation des arguments*, les individus âgés de 18 à 25 ans ($M = 5,35$) ont obtenu un score supérieur à ceux de 26 à 33 ans ($M = 5,00$, $p = 0,0021$). Les individus âgés 50 à 56 ans ($M = 6,38$, $p = 0,0006$) ont obtenu un score supérieur à ceux des autres groupes (18-25 : $p = 0$; 26-33 : $p = 0$; 34 - 41 : $p = 0$; 42-49 : $p = 0$)

En *raisonnement inductif*, les individus âgés de 26 à 33 ans ($M = 4,91$) ont obtenu un score inférieur à ceux de 18 à 25 ans ($M = 5,17$, $p = 0,0125$), 34 à 41 ans ($M = 5,15$, $p = 0,01$), 42 à 49 ans. ($M = 5,42$, $p = 0$), et 50 à 56 ans ($M = 5,92$, $p = 0$).

Analyse selon le niveau d'étude

Afin de vérifier les différences de scores entre les personnes ayant différents niveaux d'étude, un test t avec des échantillons indépendants a été réalisé pour comparer tous les facteurs du **GAAT** entre six niveaux d'étude: Niveau 1 - Diplôme de 2e ou 3e cycle universitaire / enseignement supérieur ($n = 1353$); Niveau 2 - Diplôme de 1er cycle universitaire / enseignement supérieur ($n = 314$); Niveau 3 - Diplôme d'études professionnelles ou techniques (BTS/DUT) ($n = 224$); Niveau 4 - Baccalauréat / Diplôme d'études secondaires ($n = 139$); Niveau 5 - Qualification professionnelle (CAP/BEP)($n = 99$); et niveau 6 - étudiants ($n = 59$) (voir le tableau 14).

Tableau 14. ANOVA selon le niveau d'étude.

Facteur	Niveau 1	Niveau 2	Niveau 3	Niveau 4	Niveau 5	Niveau 6	F	df	p-value
Analyse critique	5,03	4,97	5,00	4,58	5,12	5,28	2,326	5	0,000
Raisonnement déductif	5,11	4,94	5,06	4,42	5,11	5,88	6,747	5	0,018
Évaluation des arguments	5,36	5,12	5,39	4,63	5,19	5,93	6,134	5	0,041
Raisonnement inductif	5,12	5,01	5,19	4,65	5,11	5,74	4,092	5	0,001

L'ANOVA a révélé des différences significatives pour tous les facteurs. Une analyse post-hoc a été réalisée à l'aide de tests t pairwise (correction de Bonferroni) afin de voir où ces différences se produisaient.

Pour l'*Analyse critique*, les individus du niveau 4 ($M = 4,58$) ont obtenu des scores inférieurs à ceux du niveau 1 ($M = 5,03$, $p = 0,0031$), du niveau 2 ($M = 4,97$, $p = 0,0256$), du niveau 3 ($M = 5$, $p = 0,0242$), du niveau 5 ($M = 5,12$, $p = 0,0163$) et du niveau 6 ($M = 5,28$, $p = 0,008$).

Pour le *raisonnement déductif*, les individus du niveau 2 ($M = 4,94$) ont obtenu des scores supérieurs à ceux du niveau 4 ($M = 4,42$, $p = 0,0039$) et inférieur à ceux du niveau 6 ($M = 5,88$, $p = 0,0002$). Les individus du niveau 4 ($M = 4,42$) ont obtenu un score inférieur à ceux des autres niveaux : niveau 1 ($M = , p = 0$), niveau 2 ($M = 4,94$, $p = 0,0039$), niveau 3 ($M = 5,06$, $p = 0,0007$), niveau 5 ($M = 5,11$, $p = 0,0033$) et 6 ($M = 5,88$, $p = 0$).

Pour l'*évaluation des arguments*, les individus du niveau 4 ($M = 4,63$) ont obtenu des scores inférieurs à ceux du niveau 1 ($M = 5,36$, $p = 0$), du niveau 2 ($M = 5,12$, $p = 0,0111$), du niveau 3 ($M = 5,39$, $p = 0,0002$), du niveau 5 ($M = 5,19$, $p = 0,0241$) et du niveau 6 ($M = 5,93$, $p = 0$).

Pour le *raisonnement inductif*, les individus du niveau 4 ($M = 4,65$) ont obtenu un score inférieur à ceux du niveau 1 ($M = 5,36$, $p = 0$), du niveau 2 ($M = 5,01$, $p = 0,0411$), du niveau 3 ($M = 5,19$, $p = 0,0035$), du niveau 5 ($M = 5,11$, $p = 0,042$) et du niveau 6 ($M = 5,74$, $p = 0$).

3. Informations et considérations supplémentaires

Le **GAAT** a été développé principalement pour guider les décisions de ressources humaines pour le recrutement, mais son utilisation est multiple dans le milieu organisationnel. Il est important que les normes, les scores et les interprétations appropriés soient utilisés pour faciliter la sélection, mais aussi le développement de carrière, le coaching ou la gestion de la performance. Une interprétation précise des tests psychométriques aide les administrateurs et les candidats à : a) mettre l'évaluation dans un contexte; b) comparer efficacement les candidats sur les critères ciblés; c) identifier avec précision les zones à développer; d) maximiser l'adéquation entre les exigences du poste et les tendances et aspirations des candidats; et e) promouvoir l'organisation en tant que lieu de travail équitable et objectif. Le présent chapitre donne des lignes directrices aux praticiens des ressources humaines pour comprendre et interpréter de manière pertinente les scores des candidats du **GAAT** pour différents objectifs organisationnels,

Le logiciel qui génère les rapports pour **GAAT** fournit des graphiques et des commentaires personnalisés pour chacun des 4 facteurs liés au score du candidat. Le rapport ne doit être interprété que par des professionnels qui ont suivi le cours de formation qualifiante de *Central Test* et qui sont familiers avec ce manuel. La formation et la connaissance du processus d'évaluation sont essentielles car les notes doivent être interprétées à la lumière des informations concernant la fiabilité, la validité, l'erreur-type de mesure, ainsi que toute adaptation faite à l'évaluation ou séance d'évaluation. Les scores devraient également être présentés dans le contexte de l'évaluation et d'une manière claire, compréhensive et appropriée pour le public cible. Il est également important de noter que toutes les évaluations effectuées par *Central Test* ont été développées uniquement à des fins propres aux Ressources Humaines. Par conséquent, aucune conclusion clinique ne peut être tirée des résultats - pas même des scores extrêmes,

Le recrutement et la sélection constituent une fonction organisationnelle importante qui peut prendre beaucoup de temps et coûter très cher. De plus, les mauvaises décisions de recrutement ont des effets sur la productivité, les capacités et la pérennité de l'organisation. Par conséquent, il est recommandé aux recruteurs, avant d'administrer le **GAAT**, de disposer d'une description de poste précise et d'un descriptif du profil recherché détaillé, tous deux fondés sur une analyse des tâches,

En outre, le **GAAT** peut être utilisé parallèlement aux évaluations de performance, qui constituent un élément clé du développement et du succès des employés. Les évaluations de performances doivent être des évaluations justes et équilibrées des réalisations et des capacités d'un employé et doivent permettre de discuter des attentes et des réalisations,

Si le **GAAT** identifie efficacement les domaines à développer, il ne doit pas être utilisé pour suivre directement les progrès des employés. Nous recommandons donc l'utilisation d'évaluations objectives de la performance ou d'autres évaluations conjointement au **GAAT**,

Références

- Clauser, B, E., & Mazor, K, M, (1998), Using Statistical Procedures to Identify Differentially Functioning Test Items,
Educational Measurement: Issues and Practice, 17(1), 31-44,
<http://doi.org/10.1111/j.1745-3992.1998.tb00619.x>
- Dewey, J, (1910/1933), *How we think: A restatement of the reflective thinking to the educative process*, Boston: Heath,
- Drasgow, F, (1987), Study of the measurement bias of two standardized psychological tests, *Journal of Applied Psychology*, 72(1), 19-29,
<http://doi.org/10.1037/0021-9010.72.1.19>
- Fan, X,, Thompson, B,, & Wang, L, (1999), Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes, *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 56-83,
- Hanson, B, A, (1998), Uniform DIF and DIF Defined by Differences in Item Response Functions, *Journal of Educational and Behavioral Statistics*, 23(3), 244-253,
<http://doi.org/10.3102/10769986023003244>
- Holtzman, S,, & Vezzu, S, (2011), *Confirmatory Factor Analysis and Structural Equation Modeling of Noncognitive Assessments using PROC CALIS*,
<http://www.lexjansen.com/nesug/nesug11/sa/sa07.pdf>
- Hooper, D,, Coughlan, J,, & Mullen, M, (2008), Structural Equation Modelling: Guidelines for Determining Model Fit, *Electronic Journal of Business Research Methods*, 6(1), 53-60,
- Hu, L,, & Bentler, P, M, (1999), Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55,
doi.org/10.1080/10705519909540118
- Jöreskog, K, G,, & Sörbom, D, (1993), LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language,
Scientific Software International,
- McLeary, C,, & Cruise, P, (2012), Modeling Organisational Trust: Addressing Challenges of Measurement Invariance, *In Proceedings of the 11th European Conference on Research Methods (p, 470)*, Academic Conferences Limited,
- Nunnally, J, C, (1978), *Psychometric theory*, New York: McGraw-Hill,
- Pearson, K, (1895), Contributions to the Mathematical Theory of Evolution, II, Skew Variation in Homogeneous Material, *Philosophical Transactions of the Royal Society of London: Mathematical, Physical and Engineering Sciences*, 186, 343-414,

<http://doi.org/10.1098/rsta.1895.0010>

Schmidt, F, L., & Hunter, J, E, (1996), Intelligence and Job Performance: Economic and Social Implications, *Psychology Public Policy and Law*, Vol, 2, 447-472,

Trilling, B,, & Fadel, C, (2009), *21st century skills: Learning for life in our times: Learning for life in our times*, San Francisco: John Wiley & Sons,

Zumbo, B, D,, & Thomas, D, R, (1997), *A measure of effect size for a model-based approach for studying DIF*, Prince George, Canada: Edgeworth Laboratory for Quantitative Behavioral Science, University of Northern British Columbia,

NOUS CONTACTER

Service Client: **service,client@centraltest.com**

Formation et Consulting: **consulting@centraltest.com**



www.centraltest.com

© 2020 par Central Test – Empowering Talent,

Tous droits réservés,