

Le service d'admission au collège et à l'université (SACU)

par Jean-Marie JOLY *

DANS LE CADRE des travaux de la commission sur l'orientation, la consultation et la mesure dans le processus de l'admission aux études postsecondaires, on m'a demandé de décrire les méthodes employées par l'Institut de recherche pédagogique, auquel a été confiée la préparation des tests en langue française du Service d'admission au collège et à l'université (SACU), en vue d'assurer la validité de ces tests. Pour ce faire, il me sera possible de décrire des mesures qui ont effectivement été prises, alors que, dans d'autres cas, j'informerai plutôt des intentions que nous avons quant aux mesures à prendre: la validité prédictive des tests du SACU n'a pu encore être étudiée expérimentalement puisque les premiers élèves à subir ces tests n'ont pas encore terminé leur première année d'études postsecondaires.

D'autre part, à la fin de ces brefs propos, je me permettrai de soumettre deux questions qui nous occupent actuellement, dans l'espoir que vous saurez nous donner des avis sur les réponses qu'il convient d'y apporter.

* L'auteur est directeur de l'Institut de recherche pédagogique. Ce texte a été présenté à l'Assemblée générale du SACU qui se tenait à Ottawa du 18 au 20 mars 1970.

Choix des types d'items

Pour qu'un test mesure ce qu'il est censé mesurer, la première exigence, c'est que le choix des types d'items qu'on y inclut soit judicieux: il s'agit de s'assurer qu'on déclenche dans l'esprit du sujet les réactions ou les activités mentales qu'on veut avoir l'occasion d'observer. Avec les comités de préparation de tests constitués par le SACU et composés de représentants des principales régions du Canada où des élèves pourraient vouloir subir les tests en langue française, nous avons pris les décisions suivantes:

1. Pour ce qui est du *Test d'aptitude générale aux études postsecondaires* (TAGEPS), nous ne prétendons à aucune originalité. Une première distinction nous a fait établir une section verbale et une section quantitative; pour ce qui est de la première, nous employons des types d'items standards: analogies, antonymes, phrases à compléter, compréhension en lecture. Pour ce qui est de la section quantitative, nous employons des problèmes d'arithmétique, d'algèbre et de géométrie conçus de telle sorte que la difficulté, que le candidat doit résoudre, se situe au niveau du maniement des concepts quantitatifs, de

l'établissement de relations entre ces concepts, de la découverte de façons ingénieuses d'aborder un problème, mais ne présentent que très peu de difficultés de l'ordre du calcul, de l'exécution d'opérations complexes ou de connaissances particulières à l'une ou l'autre des branches des mathématiques. Cependant cette section du test contient un petit nombre d'items d'un type un peu inusité, et que nous désignons comme « suffisance des données » : il s'agit ici, pour le candidat, non pas de résoudre un problème de mathématiques, mais de porter un jugement sur la suffisance des données qui lui sont fournies pour le résoudre.

2. À l'endroit du *Test de français: langue maternelle* (TFLM) d'autre part, nous croyons avoir fait une œuvre un peu plus originale. Il s'agissait pour nous, en effet, de relever une gageure: peut-on mesurer la capacité du candidat de s'exprimer dans sa langue maternelle au moyen d'un test collectif du type « papier et crayon ». La mesure de l'expression écrite se ferait plus directement par l'utilisation d'une composition, mais l'obligation où nous nous trouvons d'élaborer un test qui puisse se corriger objectivement nous interdisait cette voie. Nous avons donc tenté, pour le choix des types d'items, de reproduire aussi exactement que possible la situation où se trouve le candidat qui doit exprimer sa pensée par écrit.

Nous avons cru pouvoir partir du postulat qu'il doit résoudre des difficultés qui se situent à au moins trois niveaux: celui du mot, celui de la phrase et celui du paragraphe. Les difficultés se situant au niveau du mot sont présentées à l'élève sous forme de phrases dans lesquelles quatre mots ou expressions ont été soulignés et identifiés de l'une des lettres A, B, C ou D. L'élève doit dire lequel de ces éléments est incorrect; s'il n'en trouve aucun, il donnera la réponse E qui signifie « pas de faute ». Bien sûr, il n'y aura jamais plus qu'un élément incorrect dans une phrase donnée.

Depuis peu, nous utilisons également l'inverse: des quatre éléments soulignés dans une phrase le candidat doit dire lequel, s'il en est, ne comporte pas d'incorrection; s'ils sont tous incorrects, il donne la réponse E.

Pour ce qui est des difficultés se situant au niveau de la phrase, on procède de la façon suivante: on présente une phrase dont une partie substantielle a été soulignée; on suggère également quatre autres façons d'exprimer l'idée contenue dans la partie soulignée. L'élève doit alors choisir l'une des cinq formulations proposées: l'originale ou l'une des quatre autres.

Enfin, nous demandons à l'élève de démontrer qu'il est apte à construire un paragraphe en procédant de la façon suivante. On présente un paragraphe, dont une phrase a été enlevée, de même que quatre, cinq

ou six phrases; à propos de chacune, il doit porter un jugement quant à sa compatibilité avec le reste du paragraphe. Il doit dire par exemple, au sujet de chacune de ces phrases, si elle convient à tous points de vue, ou si elle est inacceptable quant au sens, ou si elle est inappropriée quant au ton et au style, ou si elle comporte des erreurs grammaticales.

On aura remarqué que les deux tests, le TAGEPS et le TFLM, font usage d'items qui portent sur la connaissance de la langue.

Ce qui distingue les items basés sur la communication verbale employés dans le TAGEPS de ceux qui sont employés dans le TFLM, c'est que, dans le premier cas, l'élève a à résoudre des types de problèmes qui se présentent à lui lorsqu'il utilise sa langue à titre de « récepteur », lorsqu'il s'agit pour lui de comprendre un message: compréhension des mots utilisés, perception des relations qu'un auteur a voulu exprimer, etc. Par contre, quand il s'agit du TFLM, le candidat se trouve dans une situation où il doit utiliser la langue pour s'exprimer, comme « émetteur ».

Préparation des items

Une fois ces décisions prises, le deuxième temps consiste à recruter et à former les rédacteurs d'items. Cette tâche s'est révélée jusqu'ici assez difficile, puisqu'on ne sait pas très bien comment prédire le succès qu'une personne rencontrera dans l'exercice de cet art délicat. Résumons en disant que nous exposons aussi soigneusement que possible aux personnes choisies les exigences du travail, illustrons par certains items du type voulu qui satisfont nos critères et par d'autres qui ne les satisfont pas, démontrons comment on peut, dans un certain nombre de cas, les reviser pour qu'ils s'approchent de plus près de l'idéal, demandons à ces gens d'en produire eux-mêmes un certain nombre que nous revisons avec eux en leur indiquant, là aussi, les changements qu'il convient d'apporter. Ces mesures, aussi bien le dire, ne sont cependant pas suffisantes: la production qui nous est soumise doit ordinairement subir d'assez profondes modifications avant d'être acceptée. Ce travail est fait, d'abord par le personnel de l'Institut de recherche pédagogique (I.R.P.), puis par des experts que nous engageons à cette fin. Nous demandons ordinairement à ces

experts de répondre aux questions proposées, de sorte que toute divergence entre leurs réponses et celles que l'auteur lui-même avait indiquées nous fait voir de façon très tangible qu'il y a quelque chose qui ne tourne pas rond; de fil en aiguille, les experts finissent par nous suggérer les modifications à apporter.

Soulignons également que, pour ce qui est du test de français, tous les items sont soumis à l'Office de la langue française qui nous conseille quant à leur correction formelle, quant à l'exactitude des réponses, quant à l'importance de la difficulté soulevée, etc. Les items du TAGEPS sont également soumis à l'Office de la langue française pour correction du point de vue linguistique.

L'étape suivante consiste à soumettre les items ainsi retouchés aux comités de préparation des tests constitués par le SACU et dont mention a déjà été faite. Les experts, membres de ces comités, nous signalent les imperfections qui subsistent aux nombreuses étapes de révision et de réfection déjà mentionnées; à la lumière de ces indications, le personnel de l'I.R.P. s'efforce alors de mener l'item à un degré de perfection assez élevé pour qu'il puisse avantageusement être soumis à une épreuve expérimentale.

L'analyse des items

L'épreuve expérimentale constitue un troisième temps dans l'établissement de la validité des items. Elle consiste à administrer chacun à environ 400 élèves répartis entre le Québec et les provinces autres que le Québec, dans des proportions d'environ 90 p. 100 et 10 p. 100 respectivement, et entre les deux sexes, à part à peu près égale; les réponses de cet échantillon expérimental sont alors soumises à une analyse statistique poussée qui nous permet de déterminer si les items retenus fonctionnent comme nous l'espérons.

Cette analyse nous fournit deux indices majeurs de la qualité du fonctionnement des questions utilisées: le coefficient de corrélation bisériale entre l'item et l'ensemble des items auxquels un même élève a répondu, et un indice de difficulté. À titre d'illustration, voici les critères qui ont été employés pour la préparation des tests du SACU administrés en 1970. En premier lieu, pour être retenu, l'item devrait produire un coefficient de corrélation bisériale d'au moins .20. De plus, dans chacun de quatre sous-groupes

(garçons, filles, élèves québécois, élèves non québécois), cette corrélation devait être égale ou supérieure à .10. De plus, nous exigeons, pour retenir un item, qu'il soit de difficulté à peu près égale pour les deux sexes et pour les élèves des deux régions mentionnées. C'est ainsi, par exemple, que la différence entre l'indice de difficulté obtenu au Québec et celui qui a été obtenu hors Québec ne devrait pas dépasser 3.5. (L'indice de difficulté que nous employons est celui qui a été mis au point par le *Educational Testing Service* et désigné sous le nom de *Delta*: cet indice a une moyenne de 13 et un écart-type de 4.) Il ne doit pas non plus exister une différence trop grande quant à la difficulté pour les élèves du Québec et les élèves non québécois.

La valeur prédictive des tests

Les premiers élèves à subir les tests du SACU n'ont pas encore complété une première année d'études postsecondaires. Nous ne pouvons donc faire état d'indices statistiques descriptifs de la validité prédictive de ces tests; nous entendons bien cependant calculer de tels indices dès que les données de base auront été mises à notre disposition. Nous établirons la corrélation entre les résultats aux deux sous-tests du TAGEPS, et au TFLM, d'une part, et les résultats scolaires des élèves au cours de leur première année d'études postsecondaires. Cependant, ces mesures pouvant être raffinées, nous sommes à examiner conjointement avec le SACU et l'*Ontario Institute for Studies in Education* (O.I.S.E.), les méthodes précises qui nous donneraient les résultats les plus significatifs.

Quelques données quantitatives

Certaines données empiriques, résultant de l'administration des tests en février 1970, permettent d'observer la façon dont chacun des sous-tests inclus dans un test a fonctionné en relation avec les autres sous-tests du même test ou de la même section d'un test. Je vous présente ces données en commençant par le TAGEPS (*Test d'aptitude générale aux études postsecondaires*).

Ce test se divise en sections verbale et quantitative. Nous ne pouvons faire état de la corrélation entre ces deux sections, parce qu'elle n'a pas été produite par le programme d'analyse dont nous faisons actuellement usage. La section verbale, cependant, est composée de deux types majeurs d'items: détachés

(analogies, antonymes, phrases à compléter) et reliés entre eux, c'est-à-dire ceux que nous employons pour mesurer la compréhension en lecture. Les données qui suivent reflètent cette division.

TABLEAU I
Intercorrélations du TAGEPS, section verbale

	items détachés	compréhension	total
items détachés	1.00	.70	.86
compréhension	.70	1.00	.97

Pour ce qui est de la section quantitative du TAGEPS, là encore nous divisons les items en deux: « réguliers », c'est-à-dire ceux qui présentent des problèmes de type traditionnel et portant sur l'arithmétique, l'algèbre ou la géométrie, et ceux que nous nommons « suffisance des données », où l'élève n'a qu'à se prononcer sur le fait que, avec les données présentées, il peut ou ne peut pas résoudre le problème proposé. Voici les données quantitatives résultant de cette division.

TABLEAU II
Intercorrélations du TAGEPS, section quantitative

	items réguliers	suffisance des données	total
items réguliers	1.00	.59	.90
suffisance des données	.59	1.00	.88

Le test TFLM (*Test de français: langue maternelle*) se divise en trois sections: usage des mots, construction de phrases et élaboration de paragraphes. L'administration du test en février 1970 fournit les intercorrélations suivantes.

TABLEAU III
Intercorrélations du TFLM

	mots	phrases	paragraphes	total
mots	1.00	.56	.54	.83
phrases	.56	1.00	.50	.82
paragraphes	.54	.50	1.00	.85

Je tiens à signaler, en relation avec les trois tableaux qui précèdent, un point important: les corrélations rapportées entre un sous-test et l'ensemble du test ne doivent pas être considérées comme finales. En effet, le programme d'analyse d'items dont nous faisons usage comporte l'inconvénient de ne pas tenir

compte du fait que la note à un sous-test contribue à la cote totale; la corrélation test-sous-test que nous obtenons est donc artificiellement gonflée. Le temps ne nous a pas permis de calculer des coefficients de corrélation plus justes; nous le ferons dès que possible.

Conclusion

On a souvent fait mention que, pour un bon nombre d'élèves, les tests du SACU ne sont pas d'une utilité très grande: il s'agit de ceux qui, tout au long de l'école secondaire, ont obtenu des résultats honorables, et qui seront admis dans une institution post-secondaire quels que soient leurs résultats aux tests du SACU. Je partage partiellement cet avis: il y aurait lieu de le compléter en signalant qu'éventuellement tout au moins les résultats aux tests du SACU pourront servir non seulement à des fins d'admission, mais aussi à des fins d'orientation et de classification des élèves parmi les divers programmes d'études qui s'offrent à eux à leur sortie de l'école secondaire. Quoiqu'il en soit, admettons, pour fins de discussion, que les seuls élèves à subir ces tests pourraient être ceux dont le dossier scolaire n'est pas si reluisant qu'ils peuvent compter sur une admission plus ou moins automatique; il y aurait alors lieu de se demander si ces seuls élèves ne devraient pas être astreints à produire des scores. Le problème qui se poserait à ce moment, serait l'établissement de normes. Puisqu'en principe sont astreints à subir ces tests tous les candidats à l'admission dans un établissement post-secondaire qui a décidé d'en faire usage, les normes s'établissent simplement en se basant sur les cotes obtenues par cette population, et l'interprétation des résultats d'un candidat est relativement facile. Il n'en serait pas de même si les élèves se présentaient aux tests librement, et seulement s'ils croient que les résultats seraient de nature à faciliter leur admission. Comment conviendrait-il alors de procéder pour que les résultats de ces élèves soient quand même interprétables par les établissements où ils demandent à être admis? Puisqu'on peut croire qu'en général les élèves qui se présenteraient aux tests constitueraient la moitié inférieure de la distribution des talents, comment les institutions qui reçoivent ces candidats pourraient-elles être aidées dans l'interprétation des résultats?

Ma seconde question est peut-être un peu plus complexe. Le SACU administre à la fois des tests en anglais et des tests en français. Certains établissements reçoivent des demandes d'admission de candidats des deux langues et, fort honnêtement, les

autorisent à subir les tests en français ou en anglais, selon leur préférence. Le problème se pose alors de savoir si un même résultat obtenu, soit au test en français, soit au test en anglais, signifie la même chose. Or, ma conviction actuelle est la suivante: il serait sans doute possible, par des méthodes relativement simples, de déterminer si les deux séries de tests sont équivalentes en termes de difficulté; on pourrait songer, par exemple, à recruter un échantillon de sujets dont le bilinguisme aurait été démontré par les mesures appropriées, faire subir à ces personnes l'une et l'autre séries, et déterminer la relation qui existe entre les deux distributions de notes. Mais cela ne résoudrait pas le vrai problème; il faudrait également tenir compte du fait que l'élève qui, en

vue d'être admis dans une institution anglophone par exemple, a préféré subir les tests en français, pourrait rencontrer des difficultés au niveau de l'utilisation de la langue anglaise pour fins d'étude. En d'autres mots, il me semble possible qu'un élève extrêmement brillant réussisse admirablement les tests en français, mais ne réussisse pas très bien des études en anglais parce qu'il ne maîtrise pas assez bien la langue de Shakespeare. Il conviendrait d'établir, à la suite d'observations empiriques, une équation de régression multiple dans laquelle entreraient, non seulement les tests du SACU, mais aussi une mesure de compétence dans la langue où l'enseignement sera reçu et, peut-être aussi, une mesure d'aptitude à l'apprendre rapidement.