

Quels sont les apports de la psychologie cognitive pour articuler évaluation et apprentissage ?¹



DANIEL GAONAC'H

UNIVERSITÉ DE POITIERS (FRANCE)

Introduction

L'accent mis sur l'intérêt des évaluations formatives repose sur l'idée que le fait d'évaluer doit permettre (et peut permettre) d'améliorer la qualité des apprentissages réalisés par les élèves. Les écrits qui visent à théoriser ce concept (par exemple Allal, 1991) mettent souvent en avant l'idée qu'un tel mode d'évaluation vise à une *régulation* efficiente de l'enseignement et/ou des apprentissages qui en découlent. Ce qui caractérise l'évaluation formative n'est donc pas tant l'absence de notes, de classement, de pénalisation, de publicité institutionnelle (pairs, parents). Ce mode d'évaluation implique surtout que soient fournies à l'enseignant et/ou à l'élève des données sur l'état d'un processus d'apprentissage en cours, et que ces données permettent une prise de décision sur la conduite à tenir, c'est-à-dire sur la mise en œuvre des activités supposées rendre plus efficace l'apprentissage, au regard des objectifs qui ont été fixés.

On peut ainsi retenir deux caractéristiques de cette conception :

- Même si le processus d'apprentissage inclut une séquence d'évaluation, cette séquence ne constitue pas un élément intrinsèque de ce processus ; elle en reste un élément extérieur, dont la pertinence repose sur sa capacité à modifier le déroulé du processus d'apprentissage lui-même.
- Elle semble se référer toujours à une action volontaire sur cette nouvelle séquence, à travers la prise de conscience et l'analyse du résultat de la séquence antérieure.

C'est à propos de ces deux dimensions que nous souhaitons illustrer une vision différente de l'articulation entre évaluation et apprentissage. Nous défendons ici l'idée que l'importance de cette articulation peut être mise en évidence si l'on prend en compte les processus cognitifs qui sont en jeu lorsque les situations d'apprentissage impliquent à la fois des processus de mémorisation et des processus d'évaluation. Il s'agit donc, en s'appuyant sur la connaissance qu'on peut avoir des processus d'apprentissage, d'examiner de quelle manière ces processus généraux peuvent intégrer des processus d'évaluation, et à quelles conditions cette intégration peut s'avérer particulièrement efficace. Pour étayer ce point de vue, on peut se référer notamment à deux exemples :

¹ Pour citer ce document, merci d'utiliser la référence suivante : Gaonac'h, D. (2023) Quels sont les apports de la psychologie cognitive pour articuler évaluation et apprentissage ? Dans Cnesco, *Conférence de consensus du Cnesco l'évaluation en classe, au service de l'apprentissage des élèves : Notes des experts* (pp. 148-155). Cnesco-Cnam.

- **L'effet-test**, qui montre que l'évaluation d'un apprentissage, indépendamment de toute « notation », constitue par elle-même un élément favorisant celui-ci. On examinera quelles hypothèses on peut mettre en avant pour expliquer cet effet, notamment en analysant certaines des conditions de son apparition.
- **Les recherches dans le domaine de la métacognition** qui conduisent à contester l'idée que la prise de conscience des processus en jeu dans les apprentissages soit réellement un facteur favorable à la réalisation de ceux-ci. On peut notamment mettre en avant l'existence et l'efficacité de processus de régulation implicites, à travers lesquels « l'évaluation » de son apprentissage par l'apprenant ne tient plus à la présence d'une séquence explicite le conduisant à en évaluer l'efficacité.

I. **L'effet-test : quels sont les processus impliqués ?**

Lorsqu'il y a injonction de « revoir » un cours, la stratégie spontanée est souvent de relire. Mais de nombreuses recherches montrent que la simple révision des informations à mémoriser conduit toujours à des performances inférieures, et souvent largement inférieures, à temps égal, à celles qui sont issues d'une pratique plus élaborée : il s'agit de toutes les situations où l'apprenant est conduit à retrouver par lui-même et à produire, sous différentes formes, les informations mémorisées. On désigne cet effet favorable par l'expression « effet-test ».

Cette pratique constitue un grand classique dans des domaines bien déterminés : on pense notamment à la « récitation » de textes, de poésies, souvent demandée par l'enseignant dans la classe, mais aussi pratiquée spontanément par les apprenants. Mais de nombreuses données montrent la pertinence de cette pratique dans de nombreux domaines (pour une revue, voir Roediger et Butler, 2011, et en français dans Brown, Roediger et McDaniel, 2016) : par exemple l'apprentissage d'un vocabulaire en langue étrangère (Karpicke & Roediger, 2008), l'étude d'un cours d'histoire (Carpenter, Pashler & Cepeda, 2009), ou des leçons de sciences (McDaniel *et al.*, 2011), et ce à tous les niveaux de la scolarité ou chez l'adulte.

Il est important de relever que cet effet joue y compris lorsque l'apprenant *n'a pas la possibilité de vérifier si les réponses qu'il donne sont correctes ou pas*. Certes l'effet est alors amoindri par rapport à la situation dans laquelle il y a vérification de cette exactitude (69 % de réponses correctes au test final, contre 51 % : Butler, Karpicke & Roediger, 2007), mais il existe bien par rapport à la situation contrôle sans révision sous forme de test (31 %). Si donc il existe l'effet d'une évaluation explicite, qui peut être à l'origine d'une régulation relevant d'un processus métacognitif, l'effet-test ne peut se réduire à une telle régulation. De notre point de vue, il s'agit d'un effet direct sur les processus d'apprentissage mis en œuvre dans ces situations de test. Pour comprendre la nature de ces processus, il faut s'intéresser de manière plus précise à certains aspects des conditions de réalisation de ces tests : il s'agit de **l'optimisation du délai entre apprentissage et test**, et de **la définition des contenus d'apprentissage sur lesquels porte le test**.

On relève notamment que l'effet est plus marqué si on introduit un délai entre l'étude et le test (Butler, Karpicke & Roediger, 2007) : un test immédiat n'a que très peu d'effet, et la nécessité d'un délai suffisamment long apparaît surtout lorsque la mémorisation finale est évaluée à long terme (rappel final plusieurs semaines ou plusieurs mois après les exercices). **Le délai optimal entre apprentissage et test peut dépendre de l'objet d'apprentissage**. Dans le cas d'un vocabulaire en langue étrangère,

Pyc et Rawson (2010) montrent qu'un délai de 6 minutes induit des performances meilleures qu'un délai d'une minute. Mais sur des connaissances plus complexes, on montre qu'un délai en jours voire en semaines est plus efficace. Dans le cas du cours d'histoire évoqué plus haut, les performances finales sont meilleures lorsque la révision sous forme de test est réalisée 16 semaines après la leçon initiale. Dans le cas des leçons de sciences évoquées plus haut, l'amélioration liée à l'effet-test est la plus forte si on ménage un délai d'une vingtaine de jours entre la leçon et le test, et ce même s'il s'agit d'un test unique, qui joue tout autant que 3 tests menés durant cette période. De manière générale, la répétition d'un test n'a que peu d'effet : **un seul test peut suffire à améliorer les performances, à condition qu'il soit réalisé au bon moment** (Landauer & Bjork, 1978). De plus, le caractère favorable d'un tel délai vaut y compris lorsque l'apprenant dispose d'un feedback : contrairement à ce qu'on peut penser intuitivement, le feedback le plus efficace n'est pas celui qui intervient aussitôt après que l'élève a donné sa réponse ; **il est préférable de faire un retour en fin de test**, c'est-à-dire sur l'ensemble des connaissances qui font l'objet d'un apprentissage, et non pas de manière ponctuelle sur chaque élément de ces connaissances (Mullet *et al.*, 2014).

C'est en prenant en compte ces processus différés qu'on peut comprendre en quoi des facteurs comme l'introduction d'un délai ou la prise en compte d'un ensemble d'informations reliées, peuvent constituer des conditions favorables à l'établissement de l'effet-test. **Celui-ci ne peut être alors considéré comme un simple « état des lieux » d'une séquence d'apprentissage, mais constitue en soi un processus qui s'intègre dans les processus d'apprentissage.**

II. L'explicite et l'implicite dans la régulation des apprentissages

Nous sommes partis du constat que les bénéfices d'une évaluation formative sont souvent de fait attribués à une composante métacognitive, c'est-à-dire à la conscience que l'apprenant lui-même peut avoir, spontanément ou à travers des indications qui lui sont fournies, du fonctionnement de son propre système cognitif, et dans le cas d'un apprentissage de la manière dont il met en œuvre cet apprentissage. En effet, l'opportunité de contrôler et de réguler l'avancée d'un apprentissage est supposée être liée à la connaissance par l'élève de sa progression dans cet apprentissage et des difficultés rencontrées, à une meilleure représentation par l'élève de la tâche mise en œuvre, du but des exercices proposés, et donc à la possibilité de corriger son action sur la base de cette connaissance et cette représentation.

Si de nombreuses données permettent d'affirmer que les compétences métacognitives d'un individu peuvent avoir des effets bénéfiques sur la réussite d'un apprentissage (par exemple de Boer *et al.*, 2018, pour une synthèse), il faut cependant d'abord admettre que ce n'est pas toujours le cas, et la divergence entre croyances affichées et stratégies mises en œuvre constitue un des thèmes de prédilection des recherches sur la métacognition. Pour illustrer cette divergence, on peut revenir un moment sur le dossier de l'effet-test. Alors que les données empiriques montrent de manière claire la solidité de cet effet, il est caractéristique que **les apprenants eux-mêmes sont loin d'être conscients de la supériorité de ce mode de révision**. Quand on interroge des étudiants sur la manière dont il convient d'apprendre un cours, pour une majorité d'entre eux la lecture répétée d'un chapitre est préférable à une alternance lecture-rappel (Karpicke, Butler & Roediger, 2009). Il est frappant de constater que l'absence de conscience du bénéfice tiré d'une phase de test est avérée, y compris si l'on interroge l'apprenant juste après qu'une telle phase ait été mise en œuvre. Dans leur expérience sur l'apprentissage d'un vocabulaire en langue étrangère, Karpicke & Roediger (2008) avaient montré qu'aussitôt après la phase d'apprentissage, une phase de test sur ce qui venait d'être appris était

nettement plus efficace qu'une nouvelle présentation du vocabulaire : dans un test effectué une semaine plus tard (rappel final), ils observaient 80 % de rappel correct chez les élèves qui avaient effectué un test contre 35 % chez ceux qui avaient bénéficié d'une seconde présentation. Ils avaient demandé aux participants, aussitôt après la phase de test, d'estimer le nombre de mots dont ils pourraient se souvenir par la suite : ils n'avaient alors constaté pour cette estimation aucune différence entre les deux groupes d'apprenants. Des interviews menées ensuite auprès des participants montraient que ceux qui admettaient l'utilité d'un test de ce qu'ils avaient étudié n'admettaient cette utilité qu'au titre d'un contrôle des effets de leur apprentissage, et non d'une façon de le consolider.

L'effet-test n'est pas le seul effet à propos duquel on peut évoquer une telle cécité. Un autre exemple caractéristique porte sur l'opposition apprentissage massé / apprentissage distribué : la répétition immédiate d'un matériau objet d'apprentissage, plutôt qu'un exercice différé, est le plus souvent considérée comme le moyen le plus efficace d'apprendre (Bjork, Dunlosky & Kornell, 2013). L'explication proposée est que la simple présentation (ou la lecture) répétée d'un même contenu produit chez l'apprenant une « illusion de savoir », car cette répétition accroît la facilité avec laquelle ce contenu est traité au moment de la présentation, quand bien même cette facilité ne constitue en rien un élément favorable à la mémorisation. Il s'agit là d'un phénomène très général : les prédictions faites par l'apprenant lui-même sur l'établissement d'un souvenir en mémoire sont fondées principalement sur la facilité du traitement des contenus à mémoriser. De manière plus générale encore, les apprenants ont souvent des représentations très fausses de ce qui fait qu'un enseignement est efficace pour eux (Uttl, White & Gonzalez, 2017).

Faut-il pour autant considérer que, y compris dans le cas d'une évaluation formative, la mauvaise représentation qu'on peut avoir des processus d'apprentissage réellement en jeu rende peu efficace toute prise de conscience de l'état d'un apprentissage ? On ne peut aborder ce dossier sans prendre en compte une distinction, devenue centrale dans le domaine de la métacognition, entre des **processus explicites** et des **processus implicites**. Veenman, van Hout-Wolters et Afflerbach (2006), par exemple, distinguent :

- Des processus liés à des connaissances métacognitives conscientes. Ces connaissances peuvent jouer un rôle, mais ne garantissent nullement l'efficacité d'un apprentissage, car elles peuvent être correctes ou incorrectes.
- Des processus de régulation, fondés sur des mécanismes de feedback. Ces mécanismes ne sont pas nécessairement conscients, même s'ils peuvent néanmoins déboucher, tout au moins à long terme, sur l'élaboration de connaissances conscientes.

Pour reprendre une distinction fondamentale dans les recherches actuelles sur la structure de la mémoire humaine (voir par exemple dans Gaonac'h, 2022, chapitre 1), on peut considérer que les premiers processus relèvent de connaissances déclaratives (explicites), et les seconds de connaissances procédurales (implicites, mais éventuellement explicitables). Tsalias, Sodian & Paulus (2017) ont opérationnalisé cette dimension explicite/implicite de la métacognition, chez des enfants de 10 ans et chez des adultes, à travers une épreuve d'apprentissage de paires de dessins associés. Certaines de ces paires correspondaient à des associations fréquentes (*chien – chat*), d'autres à des associations rares (*chien – tomate*). La composante explicite était évaluée à travers un questionnaire standardisé relatif aux stratégies de mémorisation. La composante implicite était évaluée à travers

deux indices : une estimation faite par chaque participant et, pour chaque paire, de sa capacité à rappeler à partir du premier dessin celui qui lui était associé ; une mesure de la durée d'étude par chaque participant de chacune des paires de dessins. On testait ensuite, après la réalisation d'une tâche distractive, le rappel de ces dessins associés. Les données recueillies conduisent à plusieurs constats :

- Il n'y a pas de lien entre la composante explicite et la composante implicite : faire preuve de davantage de connaissances métacognitives n'induit nécessairement ni des estimations ni des durées d'étude adaptées.
- Il n'y a pas de lien entre connaissances explicites et performances de rappel.
- **Le lien est en revanche avéré entre les indices de la composante implicite et ces performances.** Il est plus important chez les adultes que chez les enfants.

Ce profil de données est par ailleurs éclairé par ce qu'on observe à partir de l'évaluation des capacités exécutives des participants, celles qui permettent à l'individu la *gestion* de ses activités cognitives : attention aux éléments importants, maintien des informations pertinentes, effacement de celles qui ne le sont plus, etc. Il s'agit pour ce qui nous concerne ici de compétences de contrôle cognitif qui peuvent être en jeu notamment dans le contrôle volontaire de stratégies de mémorisation (répétition mentale de certaines informations, établissement de liens entre différentes informations, etc.). De manière apparemment paradoxale, c'est chez les enfants, et non chez les adultes, qu'on observe un lien entre capacités de contrôle et performances de rappel, ce qui peut correspondre à une utilisation plus contrôlée des stratégies de mémorisation chez les enfants. Cela ne signifie pas pour autant que ces stratégies contrôlées soient particulièrement efficaces : cela montre surtout que l'enjeu du développement cognitif, plus que sur la connaissance explicite des mécanismes mémoriels, porte sur la mise en œuvre automatique (sans contrôle) de mécanismes de régulation, ce qui relève d'un apprentissage et n'est attesté qu'assez tard chez les adolescents. Le poids plus important du contrôle cognitif chez les enfants ne tient pas à de meilleures capacités de contrôle, mais à la nécessité de compenser la plus faible automatisation de la régulation des processus d'apprentissage.

On sait depuis longtemps que l'acquisition des compétences métacognitives peut être tardive : même chez l'adulte, la mise en œuvre de stratégies efficaces n'est ni spontanée ni générale. Dès ses premières publications sur la métacognition, Flavell (1971) mettait d'ailleurs en valeur, pour rendre compte de l'acquisition de capacités métacognitives, la notion « **d'expériences métacognitives** », c'est-à-dire la confrontation répétée à des situations variées, qui constituent des occasions de prendre en compte de manière fine les situations d'apprentissage et leurs caractéristiques. Les processus de régulation implicites que nous évoquons ici tiendraient notamment à la multiplicité et à la diversité des « expériences » cognitives relatives à un contenu d'apprentissage. L'évaluation de son apprentissage par l'apprenant ne tient plus alors à la présence d'une séquence explicite lors de laquelle il en évaluerait l'efficacité, mais à la confrontation à des situations suffisamment variées dans lesquelles il est amené à mettre en œuvre des contenus d'apprentissage qui peuvent être proches mais utilisés dans des conditions différentes.

Cela revient à dire que l'acquisition de compétences métacognitives prend du temps car elle se réalise au fur et à mesure qu'un individu est confronté à une variété de situations dans lesquelles il est amené à (re)mobiliser des apprentissages. De ce point de vue, l'adulte a plus de compétences métacognitives parce qu'il a été confronté à davantage de situations, alors même que la mise en œuvre de ces

compétences n'apparaît pas nécessairement de manière explicite lors de la réalisation d'un apprentissage.

On peut ainsi aboutir à la conclusion que **la fréquence, et surtout la variété, des situations d'évaluation, constituent en soi des conditions favorables au développement de compétences de régulation qui ne relèvent pas nécessairement d'une régulation explicite**, liée à la prise de conscience de l'état d'un apprentissage. Les situations d'évaluation doivent aussi permettre à l'apprenant d'être confronté, à travers différents contenus d'apprentissage, à des contraintes d'apprentissage suffisamment variées pour que les mécanismes implicites, dont on connaît maintenant la puissance, puissent devenir opérationnels. On pourra rapprocher ce point de vue de celui qui a été largement développé par tout un pan de la psychologie cognitive consacré à « l'expertise », dont on a montré, notamment à l'origine à propos du jeu d'échecs, qu'elle pouvait relever fortement de la référence implicite et automatisée à des situations prototypiques surentraînées (Gaonac'h, 2019, chapitre 8).

Cette référence aux dimensions implicites du fonctionnement cognitif conduit à insister sur le fait que la mise en œuvre, dans les apprentissages, de stratégies explicites, y compris sans doute dans le cadre d'une évaluation formative, est loin d'être spontanée et efficace. Si l'on suit le raisonnement qui prévaut dans la mise en exergue de processus implicites de régulation, c'est sans doute à travers la confrontation répétée à des contextes variés d'évaluation que peuvent être construites des compétences de régulation qui peuvent rester implicites sans perdre leur efficacité : c'est alors, **plus que leur explicitation, la multiplicité et la diversité des expériences cognitives relatives à un contenu d'apprentissage qui contribuent à cette efficacité.**

On peut rapprocher ce point de vue de celui défendu notamment par Lieury (1991 ; voir Gaonach, 2022, Chapitre 1, pour une synthèse), qui argumente que la construction des connaissances doit s'appuyer sur la présentation d'« épisodes » variés. Ce point de vue peut notamment être illustré simplement à travers les recherches sur le rôle des exemples dans les apprentissages. Les recherches sur l'utilisation des exemples démontrent bien leur pertinence pédagogique, mais concluent souvent à la nécessité d'introduire, pour un même exposé, plusieurs exemples présentant des caractéristiques différentes. Cet effet est particulièrement marqué pour les débutants ou les élèves en difficulté. Chi *et al.* (1989) ont ainsi analysé la façon dont des étudiants résolvent des problèmes de mécanique lorsqu'ils doivent transférer un entraînement initial dans des situations variées. Ils constatent que les étudiants ayant un faible niveau dans la discipline doivent retourner aux exemples fournis lors de l'entraînement, alors que pour les étudiants de bon niveau, la référence aux exemples pertinents étudiés auparavant est instantanée. Ainsi, la différence entre les étudiants de différents niveaux ne tient pas tant à des capacités de raisonnement qu'à une exploitation automatique d'exemples pertinents, rendue possible par la confrontation répétée à des exemples suffisamment variés.

Conclusion

Nous avons à travers cet exposé cherché à illustrer deux principes qui nous paraissent éclairer en quoi une évaluation formative peut constituer un déterminant important des apprentissages, non pas comme un « contrôle » extérieur au processus d'apprentissage, mais comme une composante inhérente à ce processus. **L'évaluation doit être conçue comme un élément intégré dans le processus d'apprentissage** en conformité avec ce qu'on connaît des processus de consolidation et de récupération des informations en mémoire, et de la manière dont les apprenants peuvent exercer un contrôle sur ces processus.

Références

- Allal, L. (1991). Vers une pratique de l'évaluation formative : matériel de formation continue des enseignants. De Boeck.
- Bjork, R.A., Dunlosky, J. & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417-444. <https://doi.org/10.1146/annurev-psych-113011-143823>
- Bjork, E.L., Little, J.L. & Storm, B.C. (2014). Multiple-choice testing as a desirable difficulty in the classroom. *Journal of Applied Research in Memory and Cognition*, 3(3), 165-170. <https://doi.org/10.1016/j.jarmac.2014.03.002>
- Brown, P.C., Roediger, H.L. & McDaniel, M.A. (2016). Mets-toi ça dans la tête : Les stratégies d'apprentissage à la lumière des sciences cognitives. Markus Haller.
- Butler, A.C., Karpicke, J.D., & Roediger III, H.L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, 13(4), 273-281. <https://doi.org/10.1037/1076-898x.13.4.273>
- Carpenter, S.K., Pashler, H. & Cepeda, N.J. (2009). Using tests to enhance 8th grade students' retention of U.S. History facts. *Applied Cognitive Psychology*, 23(6), 760-771. <https://doi.org/10.1002/acp.1507>
- de Boer, H., Donker, A. S., Kostons, D. D. & van der Werf, G. P. (2018). Long-term effects of metacognitive strategy instruction on student academic performance: A meta-analysis. *Educational Research Review*, 24, 98-115.
- Flavell, J.H. (1971). What is memory development the development of? *Human Development*, 14(4), 272-278. <https://doi.org/10.1159/000271221>
- Gaonac'h, D. (2019). *Quand le cerveau se cultive : Psychologie cognitive des apprentissages*. Hachette.
- Gaonac'h, D. (2022). *Les élèves et la mémoire*. Retz (Mythes et Réalités).
- Karpicke, J.D., Butler, A.C. & Roediger, H.L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory*, 17(4), 471-479. <https://doi.org/10.1080/09658210802647009>
- Karpicke, J.D. & Roediger, H.L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966-968. <https://doi.org/10.1126/science.1152408>
- Landauer, T.K. & Bjork, R.A. (1978). Optimum rehearsal patterns and name learning. In M.M. Gruneberg et al. (Eds.), *Practical aspects of memory* (pp. 625-632). Academic Press.
- McDaniel, M.A. et al. (2011). Test-enhanced in a Middle School Science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, 103(2), 399-414. <https://doi.org/10.1037/a0021782>
- Mullet, H.G. et al. (2014). Delaying feedback promotes transfer of knowledge despite student preferences to receive feedback immediately. *Journal of Applied Research in Memory and Cognition*, 3(3), 222-229. <https://doi.org/10.1016/j.jarmac.2014.05.001>
- Pyc, M.A. & Rawson, K.A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330(6002), 335. <https://doi.org/10.1126/science.1191465>
- Roediger, H.L. & Butler, A.C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20-27. <https://doi.org/10.1016/j.tics.2010.09.003>

Rowland, C.A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432-1463. <https://doi.org/10.1037/a0037559>

Tsalas, N., Sodian, B. & Paulus, M. (2017). Correlates of metacognitive control in 10-year old children and adults. *Metacognition Learning*, 12(3), 297-314. <https://doi.org/10.1007/s11409-016-9168-4>

Uttl, B., White, C. A. & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22-42.

Veenman, M.V.J., van Hout-Wolters, B.H.A.M. & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, 1, 3-14. <https://doi.org/10.1007/s11409-006-6893-0>

Zaromb, F.M. & Roediger, H.L. (2010). The testing effect in free recall is associated with enhanced organizational processes. *Memory & Cognition*, 38(8), 995-1008. <https://doi.org/10.3758/mc.38.8.995>