

# Questions d'actualité en docimologie : des biais cognitifs aux paramètres de l'évaluation

Conférence au Conservatoire Royal de Liège  
Dylan Dachet  
Mardi 08/03/2022



**CONSERVATOIRE ROYAL DE LIÈGE**

ÉCOLE SUPÉRIEURE DES ARTS DE LA FÉDÉRATION WALLONIE-BRUXELLES



**LIÈGE**  
université

# Au programme ...

## 1. Des biais cognitifs ...

- a. Le concept de biais
- b. Quelques biais cognitifs : expérience, définition opérationnelle, impacts, étude(s) princeps et exemple(s)

## 2. ... à leur impact sur l'évaluation

- a. L'évaluation peut-elle être neutre et objective ?
- b. L'effet *Posthumus* et l'expérience fameuse de Gjorgjevski
- c. L'évaluation est influencée par les caractéristiques des élèves
- d. L'évaluation est influencée par ses paramètres

## 3. ... en passant par le paramétrage de ses pratiques évaluatives

- a. L'évaluation, le Docteur Jekyll and Mister Hyde
- b. Des fonctions multiples, mais une visée de régulation
- c. Les paramètres de l'évaluation
- d. Le renforcement positif

## 4. Conclusion : changer l'évaluation pour changer l'école

## 5. Bibliographie

# Des biais cognitifs ...



# Le concept de biais

(Kahneman et al., 2021, p. 9-11)

Equipe A



Monde idéal

**Le biais et le bruit – l'écart  
systématique et la  
dispersion aléatoire – sont  
deux composantes  
différentes de l'erreur**

Equipe B



Biais

Equipe C



Bruit

Equipe D



Biais & Bruit



# Le concept de biais

Les biais cognitifs = les **erreurs systématiques** que notre mental a tendance à commettre lorsqu'il réfléchit (Thalemann, 2021, p. 25)

Il s'agit d'un schéma de pensée trompeur et faussement logique (Usabilis, 2018)

- Altère le jugement
- Facilite/accélère la prise de décision



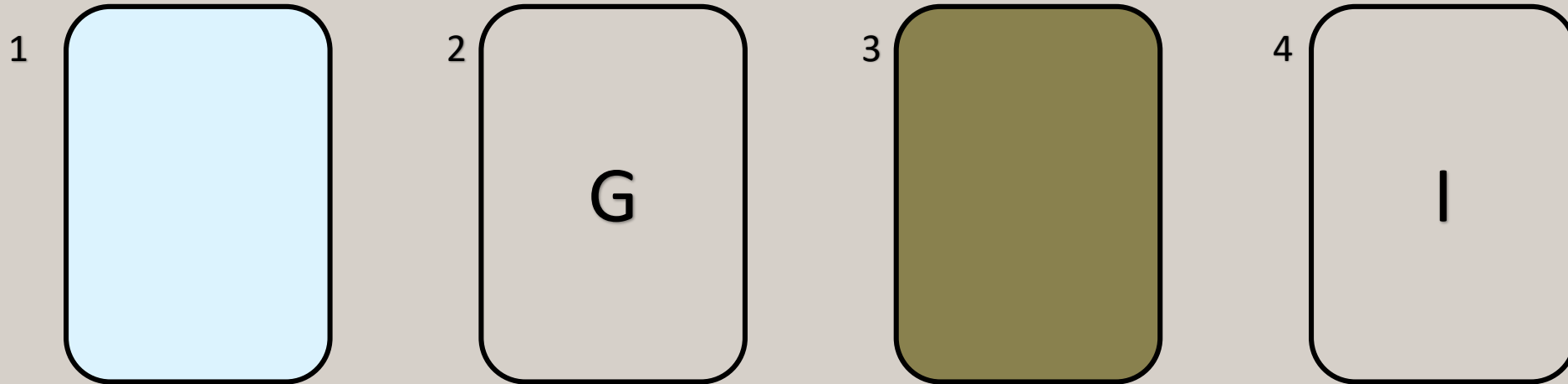
- Deux vitesses de la pensée (Kahneman, 2012) :
- (1) Système 1 : automatique et incontrôlable
  - (2) Système 2 : calculateur



# Quelques biais cognitifs

(Expérience de Wason & Shapiro, 1971)

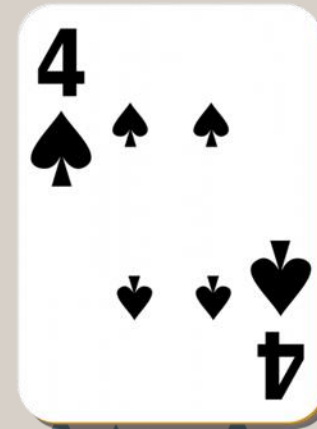
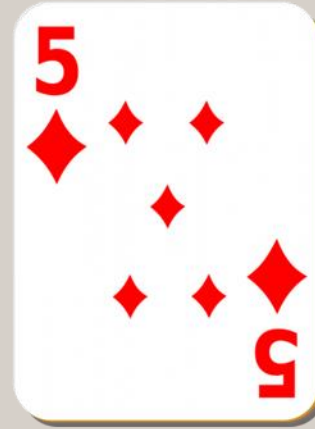
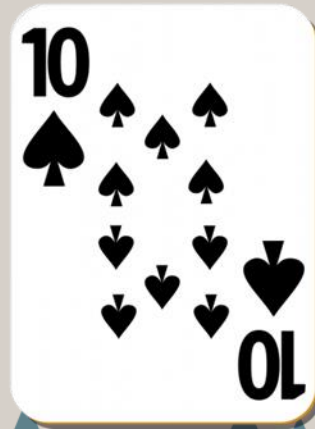
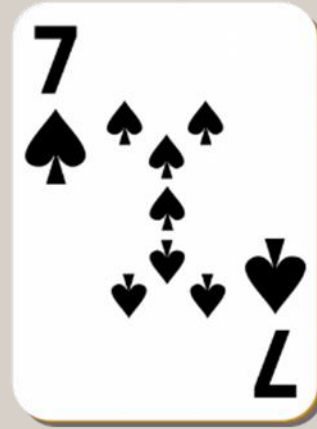
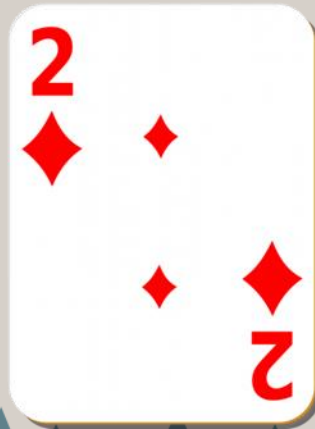
Règle : les cartes avec une consonne ont le dos foncé.



Question : lesquelles parmi les quatre cartes ci-dessus devez-vous retourner, sans en toucher inutilement, pour déterminer si la règle est valide ?

# Quelques biais cognitifs

(Tour de Magie de Thalemann, 2021)



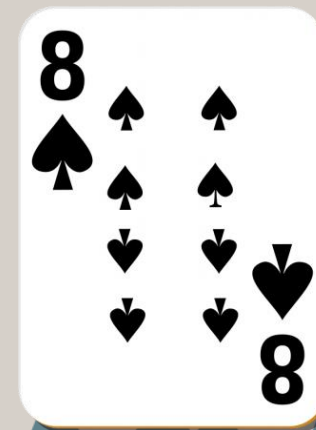
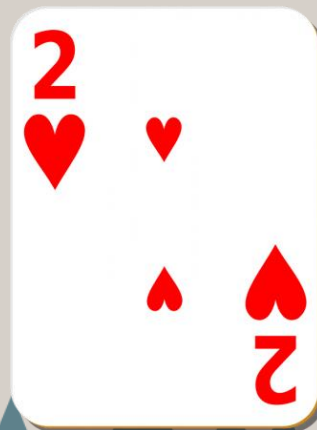
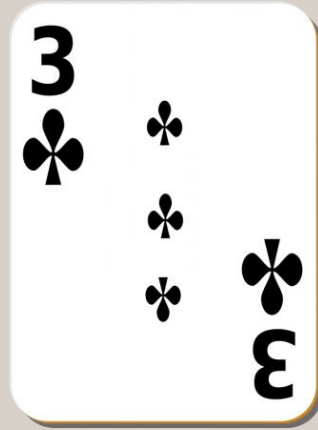
# Quelques biais cognitifs

(Tour de Magie de Thalemann, 2021)

A

# Quelques biais cognitifs

(Tour de Magie de Thalemann, 2021)



# Quelques biais cognitifs

## Le biais de confirmation

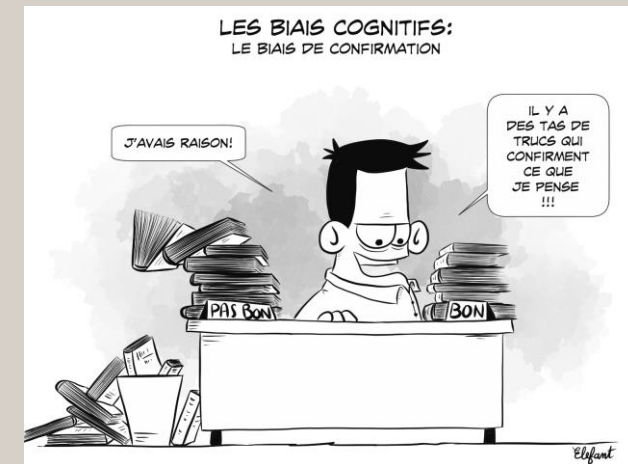
= tendance à privilégier les informations et/ou exemples qui confirment nos croyances au détriment de celles qui pourraient les remettre en question (Durand, 2019 ; 121-135 Thalemann, 2021, p. 25-26)

### Exemples :

- (1) Syndrome de l'étudiant en médecine
- (2) Étude de Lord et al. (1979) sur la peine de mort
- (3) Étude de Sunstein et al. (2017) sur le réchauffement climatique
- (4) ... Avortement, homosexualité, ...

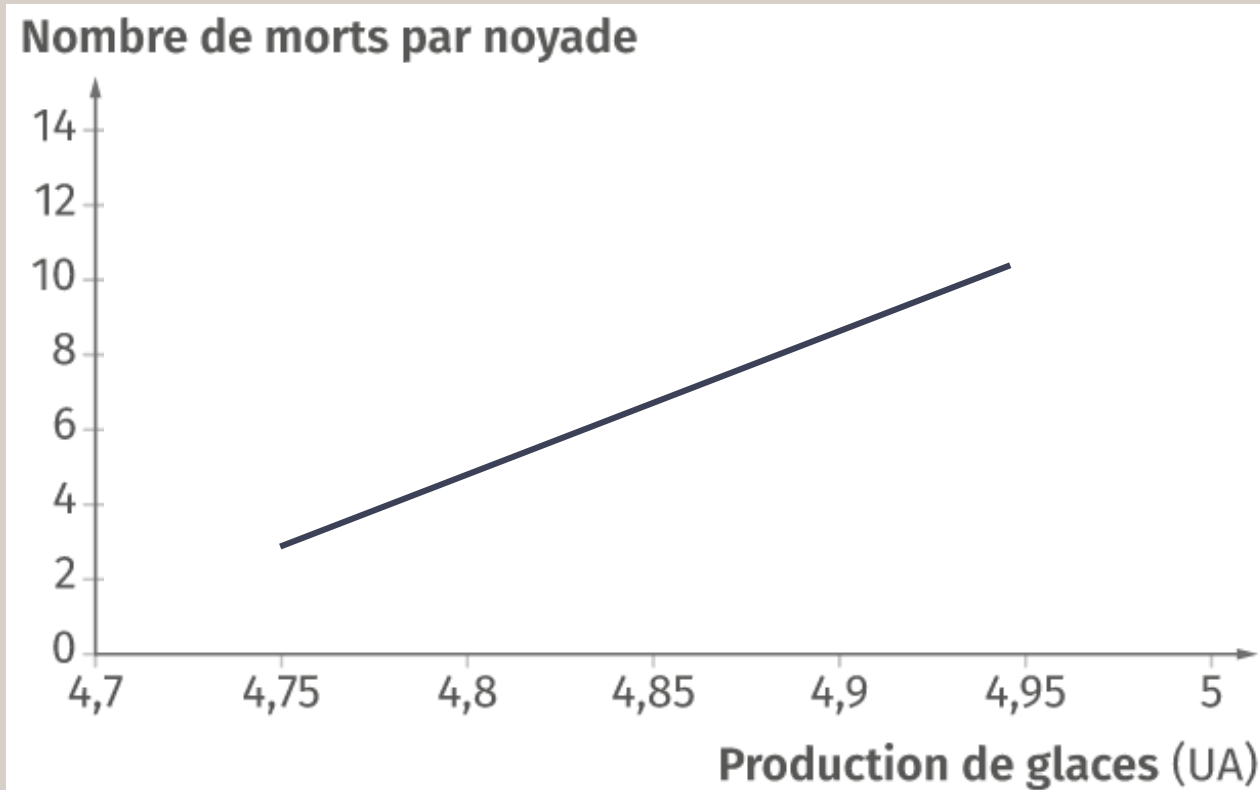
Explication : psychologie évolutionniste (Boutang & De Lara, 2009)

→ Intérêt adaptatif (Thalemann, 2021)



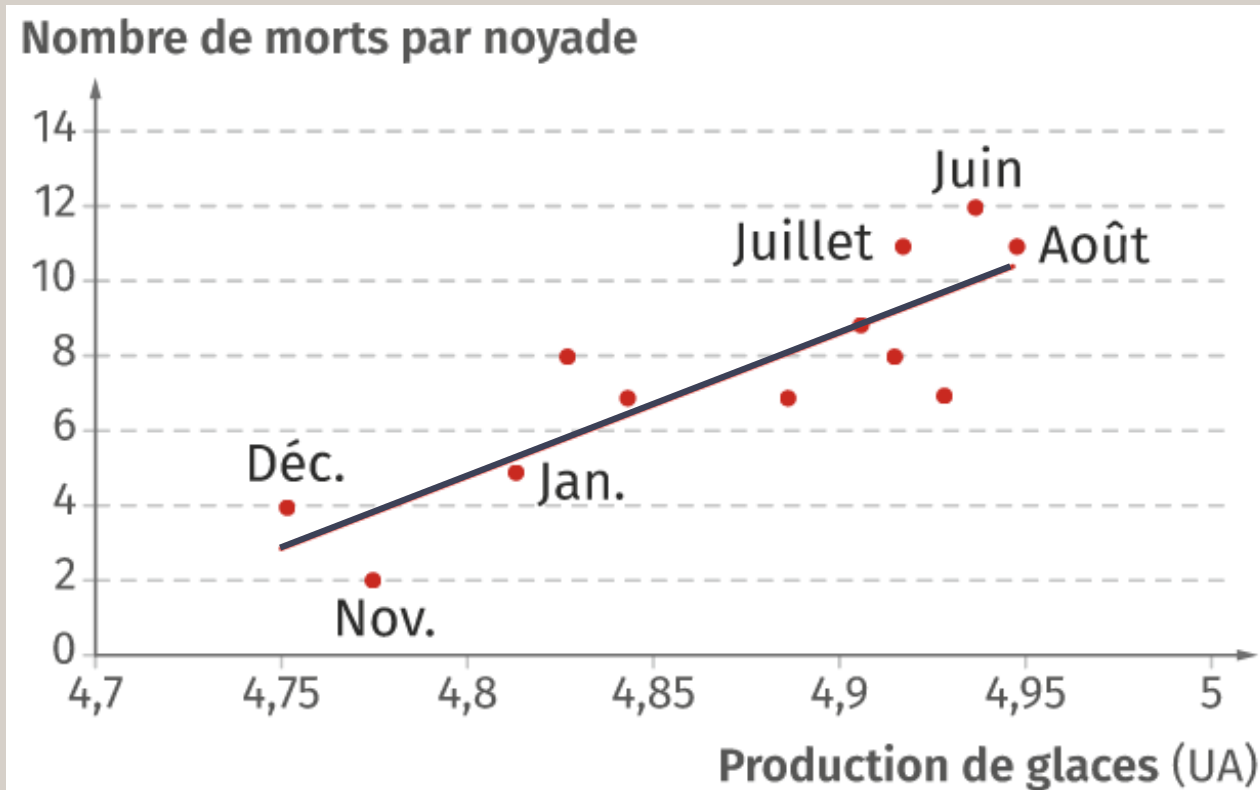
Source : [www.comicscience.net](http://www.comicscience.net)

# Quelques biais cognitifs



Source : <https://www.livrescolaire.fr/page/12138125>

# Quelques biais cognitifs



Source : <https://www.lelivrescolaire.fr/page/12138125>



# Quelques biais cognitifs

## Le biais de causalité (ou d'attribution causale)

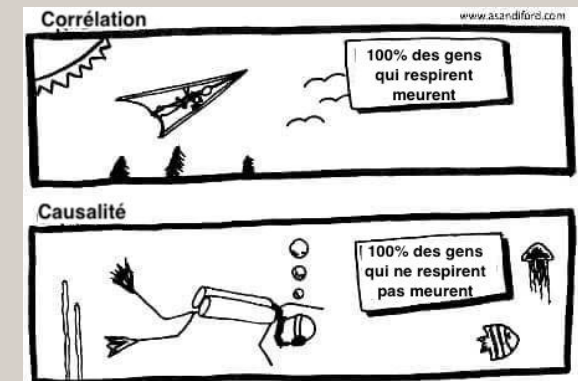
= tendance à attribuer des liens causaux entre des événements qui sont simplement corrélés et/ou dont l'occurrence simultanée est due au hasard.

Exemples :

- (1) Synchronicité : coïncidences d'événements sans lien de causalité, mais auxquelles on peut attribuer une interprétation (Durand, 2021, p, 59)
- (2) Apophénie : perception spontanée de rapport et de significations à partir de phénomènes sans aucune relation (Thalemann, 2021, p, 75)
- (3) Efficacité de nos pratiques pédagogiques (Slavin, 2007)

Explication : psychologie évolutionniste (Boutang & De Lara, 2009)

→ Intérêt adaptatif (Durand, 2021 ; Thalemann, 2021)



Source : [www.penser-critique.be](http://www.penser-critique.be)

# Quelques biais cognitifs

(Expérience de Kahneman et al., 1982)

Lucie a 31 ans. Célibataire, expressive et diplômée d'un master en droit, elle militait pour l'environnement durant ses études. Selon vous, quelle situation est la plus probable pour Lucie aujourd'hui ?

1. Lucie est professeure des écoles.
2. Lucie tient une librairie et prend des cours de karaté.
3. Lucie est assistante sociale.
4. Lucie est active dans la lutte pour le droit des animaux.
5. Lucie vend des assurances-vie.
6. Lucie est avocate.
7. Lucie est avocate et active dans la lutte pour le droit des animaux.

# Quelques biais cognitifs

## Le biais de représentativité

= tendance à considérer comme plus véridique/crédible/probable un événement correspondant à une représentation ignorante, par là, les probabilités *a priori* (Thalemann, 2021 ; Van Vyve, 2021)

### Exemples :

- (1) Erreur de conjonction : préférence à fonder un jugement sur des données personnalisantes plutôt que sur un calcul probabiliste
- (2) Probabilité qu'un objet A appartienne à une classe B : estimation en fonction du degré de représentativité de A par rapport à B



# Quelques biais cognitifs

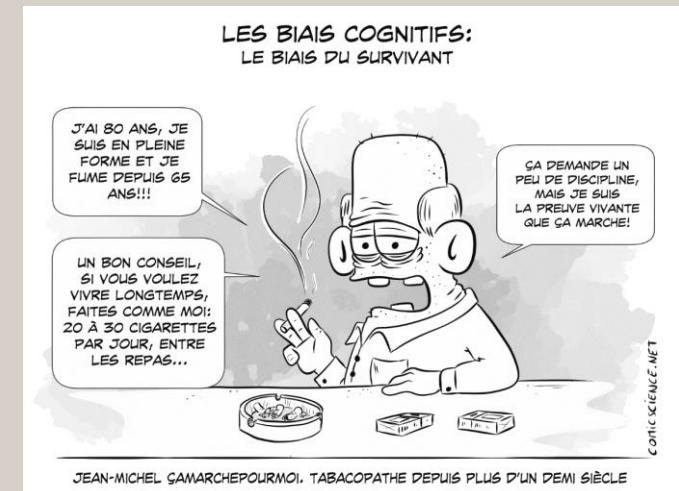
## Le biais du survivant

= tendance à surévaluer les chances de succès en se focalisation sur les succès passés sans prendre en compte leur représentativité

Exemples :

- (1) Cimetière de Diagoras (Cicéron)
- (2) Les guérissons de Lourdes
- (3) ...

	Lourdes	Hôpitaux
	67 guérissons attestés par l'Eglise	1574 rémissions spontanées
	Recensement depuis 1858	Recensement depuis 1864
	900 000 000 de visiteurs	
	300 000 000 de visiteurs (si on retire les personnes qui y vont plusieurs fois)	
	2/10 000 000	3/1 000 000 (probabilité plus importante si on inclut les rémissions du cancer)



Source : [www.comicscience.net](http://www.comicscience.net)

# Quelques biais cognitifs

(Kahneman & Tversky (1974))

- En moyenne 50 % des bébés sont des garçons.
- Une ville a deux hôpitaux.
- Dans le plus grand des deux, 45 bébés naissent chaque jour.
- Dans le plus petit, 23 bébés naissent chaque jour.
- On note sur une année le nombre de jours où plus de 60 % des bébés sont des garçons ;

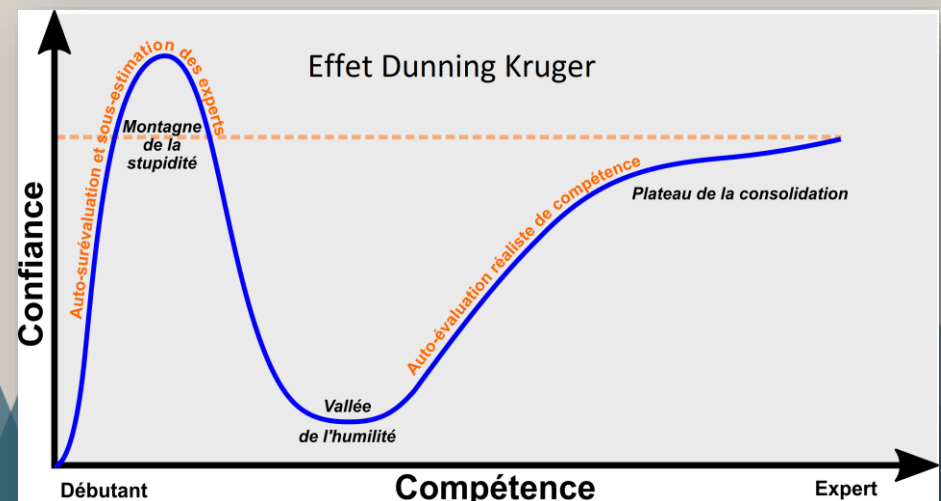
*Quel hôpital a enregistré le plus grand nombre de ces jours ?*

# Quelques biais cognitifs

**Le biais d'attente excessive d'étalement** (Thalemann, 2021, p. 75) : tendance à croire que les phénomènes aléatoires sont régulièrement dispersés, plutôt qu'apparaissant de manière regroupée (par paquets de similarité).

**Le biais d'ancrage** (Thalemann, 2021, p. 113-114) : tendance à orienter l'attention et/ou à privilégier des contenus mentaux en rapport avec un mot, une image, ... perçu(e) au préalable de façon tout à fait inconsciente → rester à la première impression.

**L'effet Dunning-Kruger** (Dunning, 2011 ; Kruger & Dunning, 1999) : tendance à surestimer ses propres compétences/connaissances, surtout dans un domaine non connu.





# Quelques biais cognitifs

**Biais mnésique** (Usubalis, 2018) : tendance à mieux mémoriser les dernières informations (effet de récence) et le première informations (effet de primauté) auxquelles on a été confrontées.

**Effet de halo** (Asch, 1946 ; Usubalis, 2018) : tendance à généraliser un jugement posé (sur une personne, une marque, ...) à partir d'un seul et uniquement élément (caractéristique personnelle, caractéristique physique, ...) ainsi qu'à sélectionner, a posteriori, les informations allant dans le sens de cette première impression

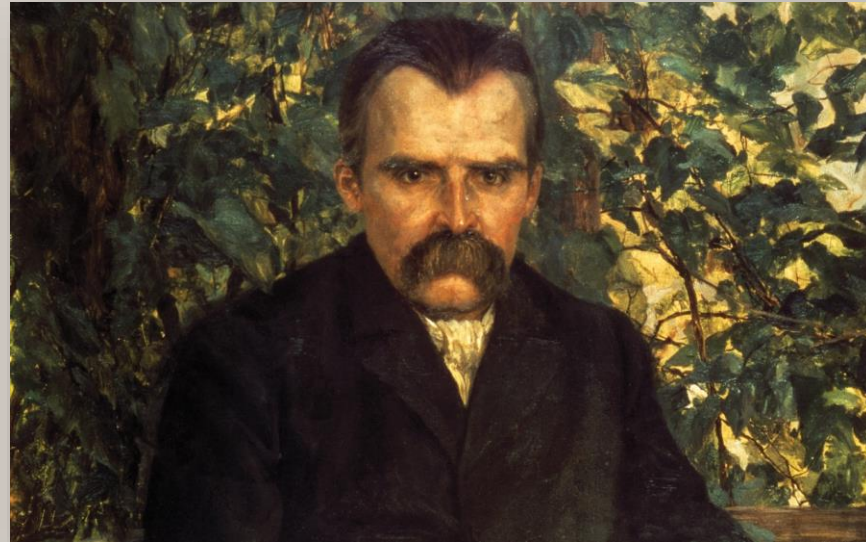
...



Et des dizaines d'autres ... Jochen Grünbeck (2021) en site 72 !!!

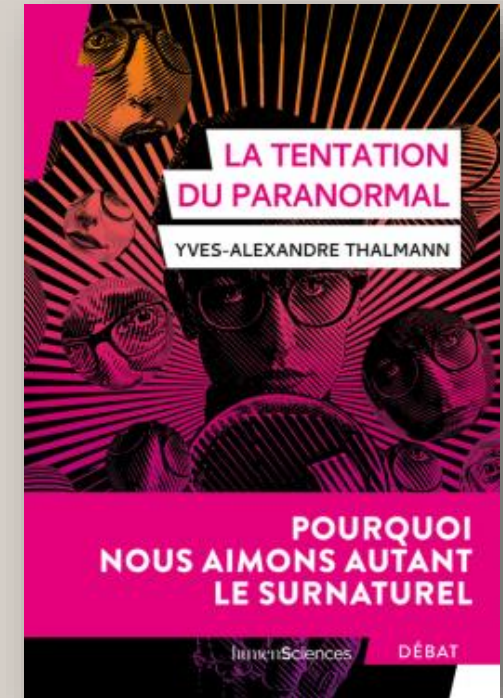
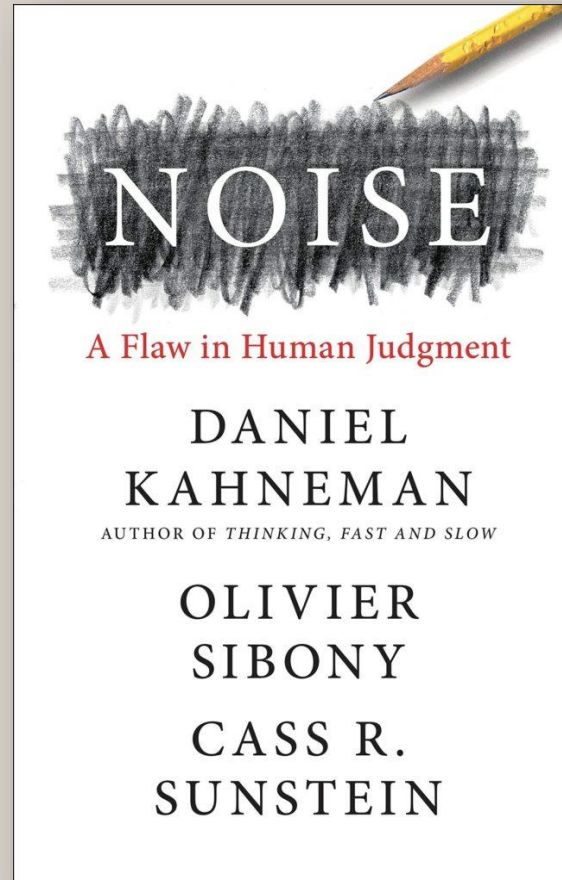
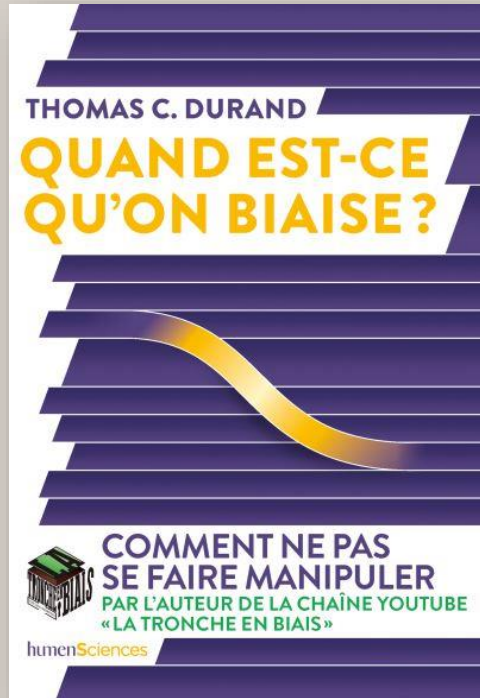
# Quelques biais cognitifs

« La croyance forte ne prouve que sa force, non la vérité de ce qu'on croit » Nietzsche (1878, « Humain, trop humain »)





# Que lire pour approfondir tout ceci ?



... à leur impact sur l'évaluation

2

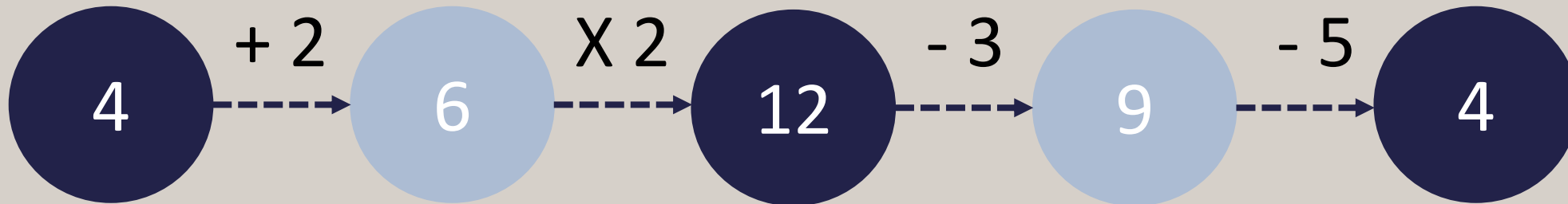
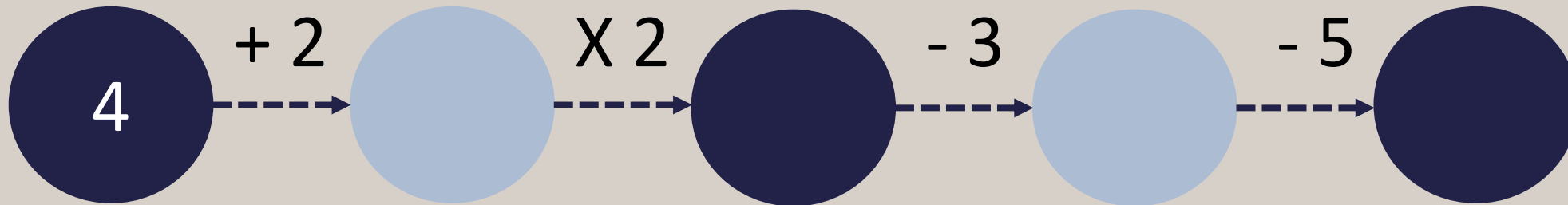
2

# L'évaluation peut-elle être neutre et objective ?



« *Évaluer consiste à **définir des critères et des indicateurs** dans le but de prendre des décisions, donc à choisir des éléments considérés comme pertinents dans **le référent** (ce qui devrait être, ce que l'on projette) et à déterminer durant l'action et à l'issue de celle-ci, si ces éléments sont bien **présents dans le référé** (ce qui est)» (Raynal & Rieunier, 2014, p. 215)*

# L'évaluation peut-elle être neutre et objective ?



Réponse de Thomas

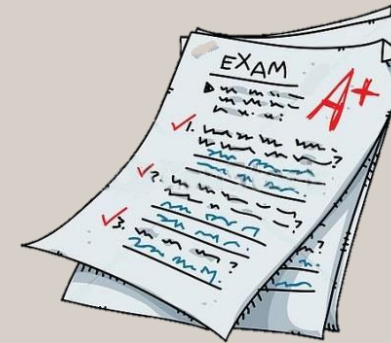


# L'effet *Posthumus* et l'expérience fameuse de Gjorgjevski

(cité par Rot et Butas, 1959)

## Étape 1

5 correcteurs

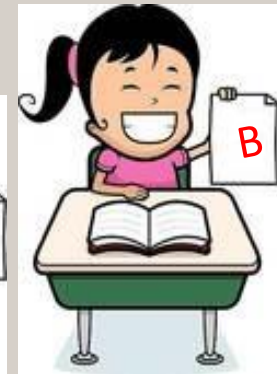
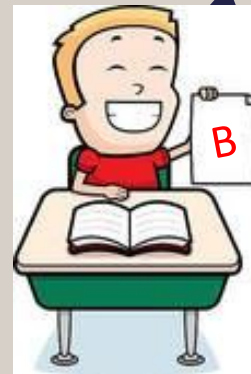


100 épreuves

## Étape 2

15 copies notées « Bien »

4 autres correcteurs

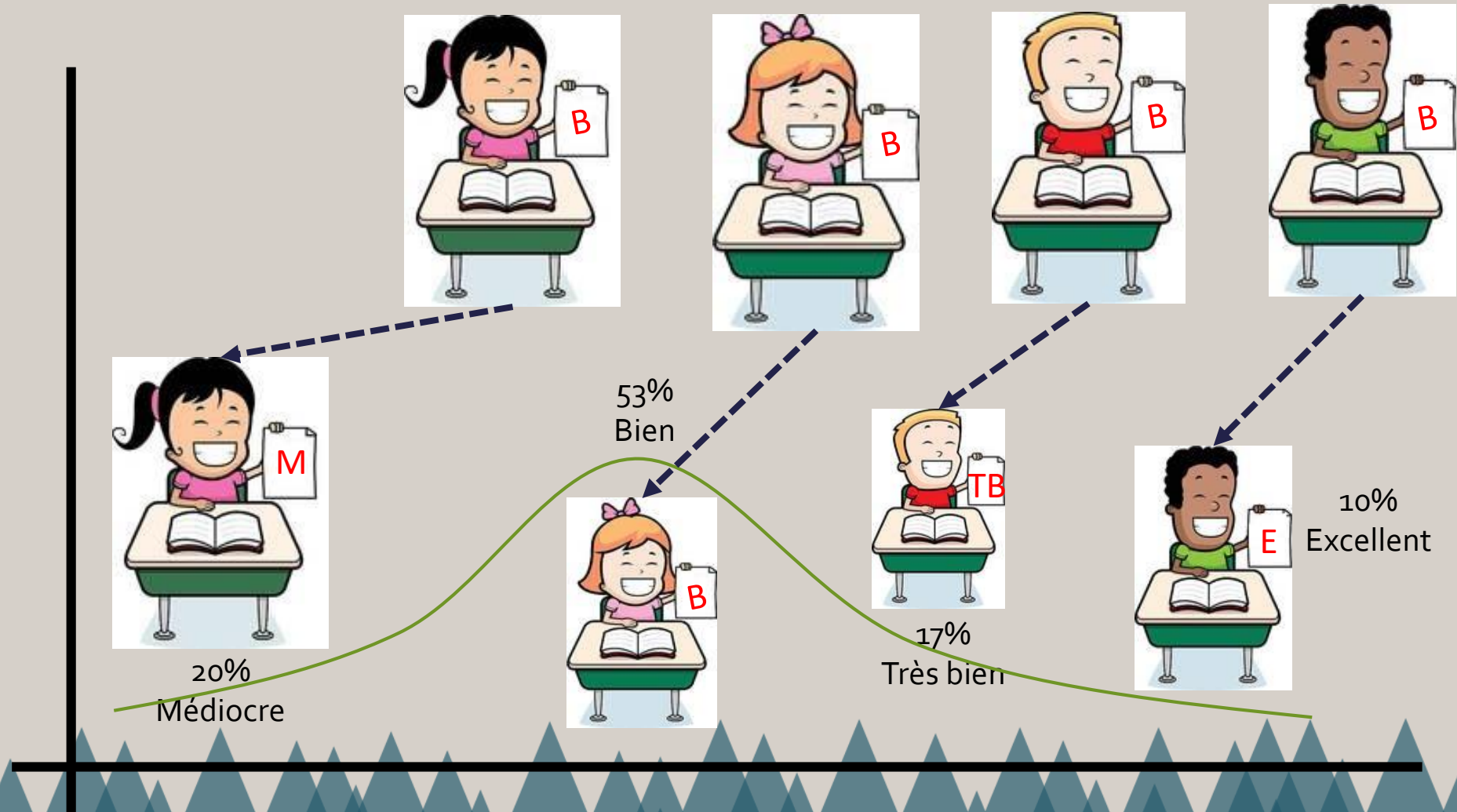




# L'effet *Posthumus* et l'expérience fameuse de Gjorgjevski

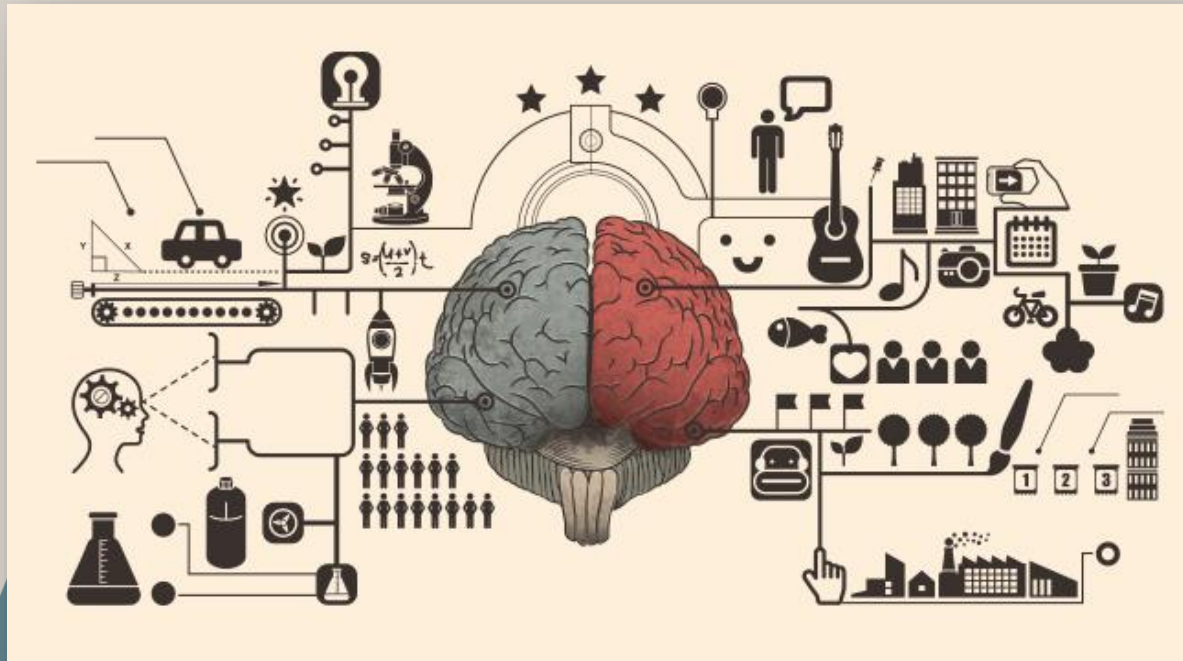
(cité par Rot et Butas, 1959)

## Résultats

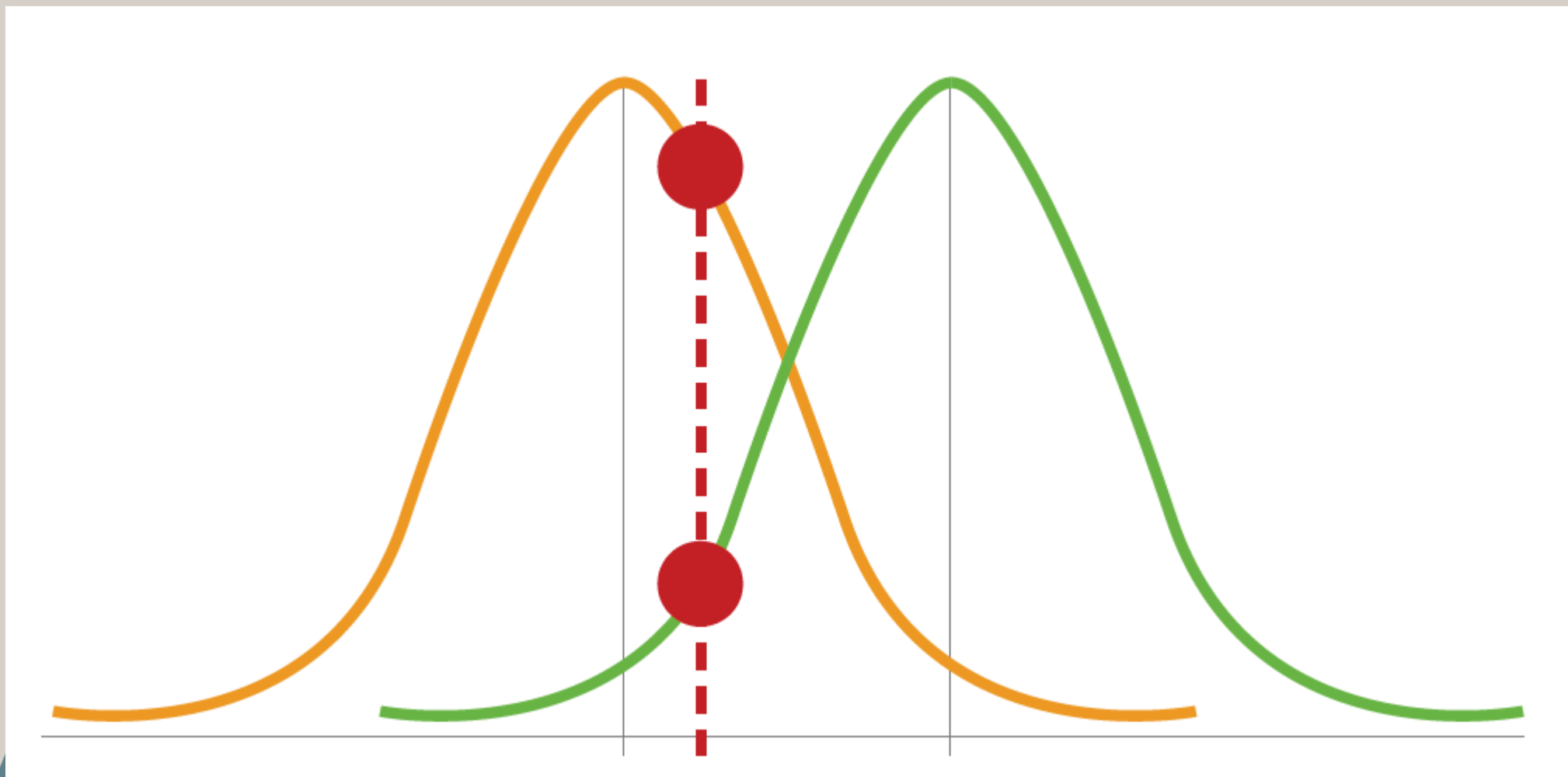


# L'effet *Posthumus* et l'expérience fameuse de Gjorgjevski

« *L'enseignant adapte le niveau de son enseignement et ses appréciations des performances des élèves de façon à conserver d'année en année, approximativement la même distribution (gaussienne) de notes* » (De Landsheere, 1980)



# L'effet *Posthumus* et l'expérience fameuse de Gjorgjevski





# L'évaluation et les caractéristiques des élèves

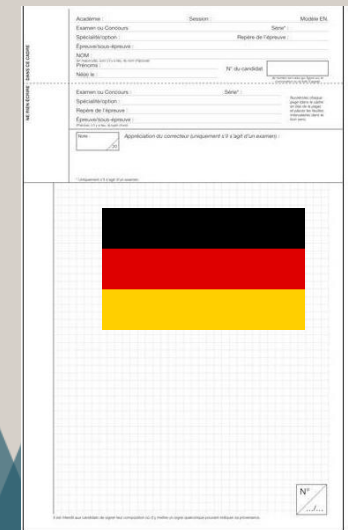
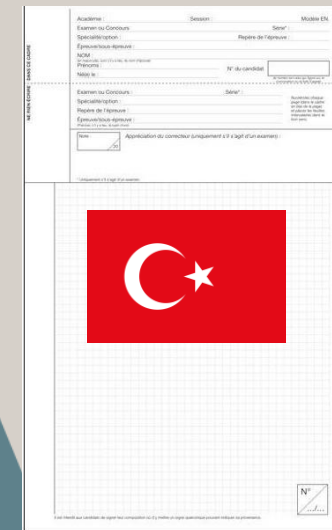
(Sprietsma, 2013)

L'évaluation des performances des élèves par des enseignants peut-elle être biaisée par leur connaissance du milieu social d'un élève ?

Etude expérimentale :

- (1) Recueil de dissertations rédigées par deux classes anonymes
- (2) Sélection aléatoire de 10 dissertations
- (3) Envoie de la série chaque enseignant avec assignation aléatoire d'un prénom à chaque copie
  - Des prénoms à consonance germanophone (Max, Stefan, ...)
  - Des prénoms à consonance turque (Hakan, Gönül, ...)
  - (une des communautés ethniques les plus représentées)

88 enseignants du primaire (58 écoles différentes)



# L'évaluation et les caractéristiques des élèves

(Rangvid, 2015)

L'évaluation des performances des élèves par des enseignants peut-elle être biaisée par leur connaissance du milieu social d'un élève ?

Comparaison entre : (1) les données issues des examens étatiques externes et (2) les notes des enseignants

**Table 1**  
Descriptive statistics by subsamples.

Exam score=7		Boys	Girls	Low parental education	High parental education	Migrants	Natives	All students
Teacher scores	Mean	6.61	7.09	6.40	7.37	6.51	6.89	6.86
	SD	2.13	2.02	2.14	1.99	2.21	2.08	2.09
Sample size	No. scores	762,558	838,901	387,618	335,080	103,127	1,498,332	1,601,459

*Pour les élèves qui ont obtenu un score identique de 7 au test étatique, les notes des enseignants étaient plus hétérogènes : elles variaient de 6,40 pour les élèves dont les parents ont un faible niveau d'éducation à 7,37 pour les élèves dont les parents ont un haut niveau d'éducation.*

# L'évaluation et les caractéristiques des élèves

(Lafontaine & Monseur, 2009)

L'évaluation des performances des élèves par des enseignants peut-elle être biaisée par leur connaissance du genre d'un élève ?

Etude expérimentale :

(1) Copies fictives construites par un professeur de mathématiques : copie très faible, une copie assez faible, une copie moyenne et une bonne copie

(2) Seule information disponible

→ prénom clairement masculin (Bernard et Nicolas)

ou féminin (Chloé et Emilie)

48 professeurs de mathématiques

	Filles	Garçons
Moyenne	5,250	5,375

	C. très faible	C. faible	C. moyenne	Bonne C.
Moyenne filles	2,042	5,250	6,354	7,354
Moyenne garçons	1,750	4,917	6,937	7,895

# L'évaluation et ses paramètres

(Schwarz, 1999)

La manière de poser une question influence les réponses obtenues :

1. A quel point avez-vous réussi dans la vie ?

0	1	2	3	4	5	6	7	8	9	10
-5	-4	-3	-2	-1	0	1	2	3	4	5



2. A quelle fréquence apparaissent les symptômes

<b>Deux fois par mois ou moins</b>	<b>Plus de deux fois par mois</b>	<b>Plus de deux fois par semaine</b>	<b>Plusieurs fois par jour</b>	62 %
<b>Jamais</b>	<b>Une fois par an</b>	<b>Une fois par mois</b>	<b>Plus de deux fois par mois</b>	39 %



3. Quelle est la chose la plus importante à enseigner aux enfants ?

- 61 % cochent la réponse : penser par eux-mêmes
- 5 % le mentionnent spontanément lors d'une question ouverte.

# L'évaluation et ses paramètres



# L'évaluation et ses paramètres

	Évaluation normative 	Évaluation formative 
<b>Définition</b>	utilise des indicateurs tels que des côtes, lettres ou jugements de valeur qui servent parfaitement l'objectif de comparaison avec une norme et entre individus.	fournit un feedback (ici qualitatif) spécifique et détaillé
<b>Visée</b>	comparer les performances de la personne évaluée à celles d'autres personnes	adapter les activités d'enseignement et d'apprentissage et fournir des commentaires pertinents
<b>Contexte</b>	Méritocratique	Correctif/de Maîtrise
<b>Effet</b>	Notation d'un test attribué à un SES - ou SES + Étudiants inscrits dans un programme sélectif SSE + > SSE - (Batruch, Autin & Butera, 2017)	Notation d'un test attribué à un SES - ou SES + Étudiants inscrits dans un programme non-sélectif SSE + = SSE - (Batruch, Autin & Butera, 2017)



# L'évaluation et ses paramètres



**VS**

**Dylan**      **Charles**



**11 fautes claires**  
**6 fautes ambiguës**

Les expériences

**N = 196**  
**à 374**

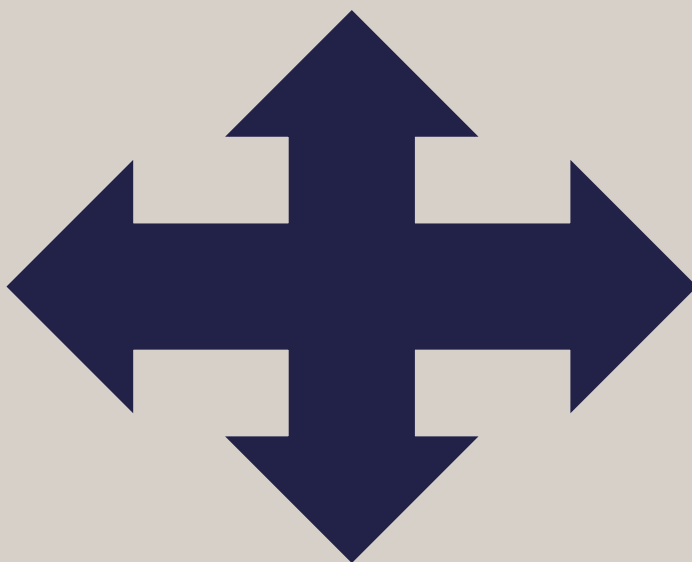


# L'évaluation et ses paramètres

**Dylan**



Les expériences



**Charles**



**Dylan**



**Charles**





# L'évaluation et ses paramètres

(Autin et al., 2019)

## Expérience 1

Etude expérimentale:

- Double manipulation :
  - Note (normatif) vs. commentaire(s) (formatif)
  - SES faible vs. fort (via des indicateurs tels que le prénom, le statut professionnel des parents, activités extracurriculaires, ...)
- N = 196 étudiants (imaginant qu'ils sont des enseignants)
- Correction d'une dictée (1-6) contenant 11 fautes claires et 6 fautes ambiguës

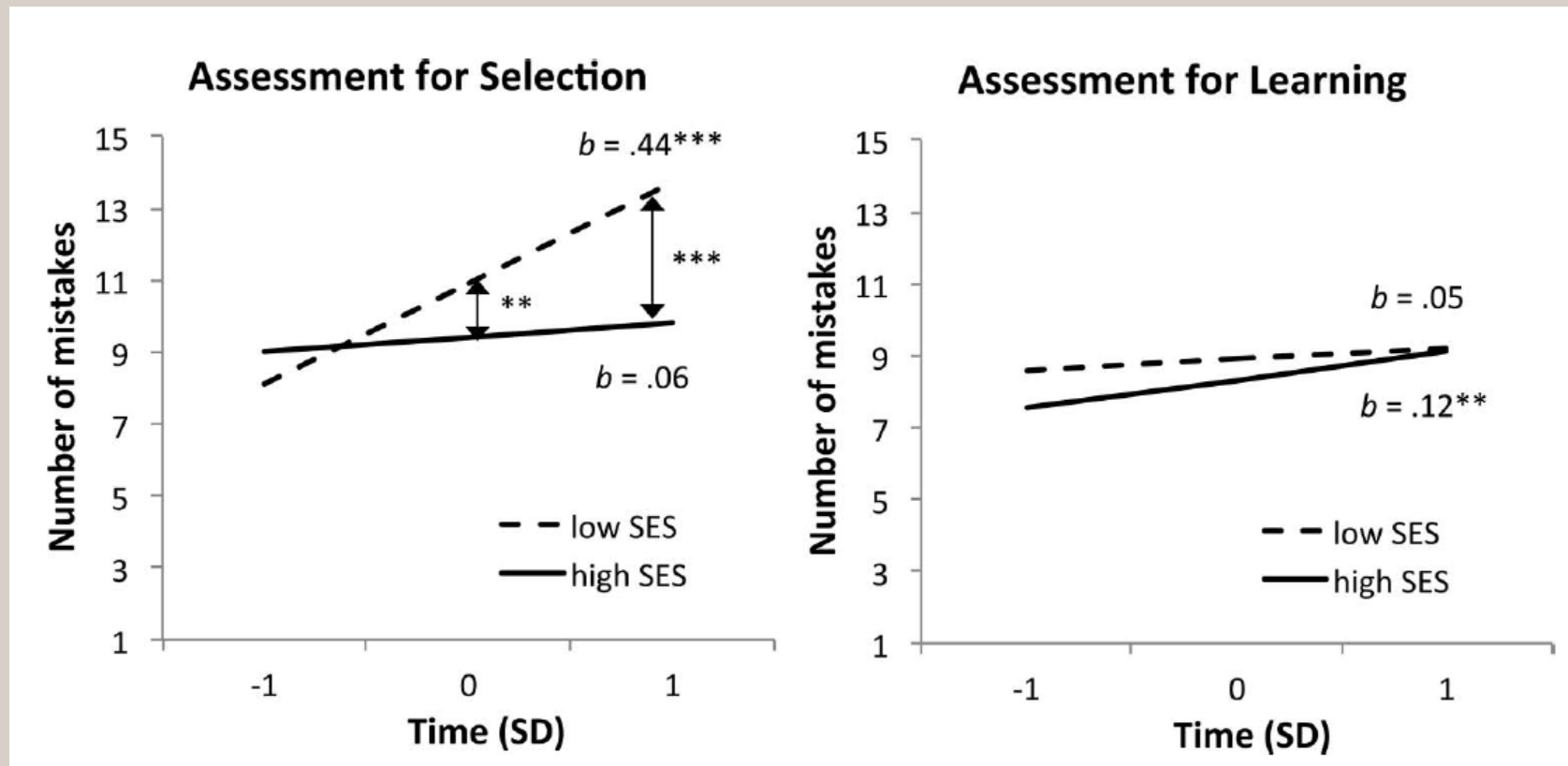
Variables étudiées :

- Temps de correction → plus grand dans le groupe « commentaire » que dans le groupe « Note »
- Nombre de fautes repérées :
  - Note (M = 10.16, SD = 2.53) > Commentaire(s) (M = 8.61, SD = 2.18).
  - SES faible (M = 9.90, SD = 2.54) > SES fort (M = 8.87, SD = 2.16)
- Effet positif du temps sur le nombre de fautes repérées (b = 0.17).

# L'évaluation et ses paramètres

(Autin et al., 2019)

## Expérience 1



# L'évaluation et ses paramètres

(Autin et al., 2019)

Estimation de l'ampleur de l'effet de la modulation de l'écart de performance dû au SSE de l'élève en fonction de l'orientation de l'évaluation (sélective vs. éducative):

$$D = 0.19$$

$$p = 0.002$$

« Cette méta-analyse interne prouve que **les évaluateurs créent artificiellement un plus grand écart de performance SSE lorsque l'évaluation est utilisée pour sélectionner plutôt que pour favoriser l'apprentissage**. L'ampleur de l'effet est faible, mais nous pensons néanmoins qu'elle doit être interprétée à la lumière de la durée de l'éducation et de la fréquence de l'évaluation. De très petites différences dans les évaluations répétées peuvent avoir des conséquences importantes sur les expériences globales et les résultats scolaires des étudiants lorsqu'elles s'accumulent au fil du temps » (Autun, Batruch & Butera, 2019, p. 729).

... en passant par le paramétrage de ses  
pratiques évaluatives

3

3



# L'évaluation, le Docteur Jekyll and Mister Hyde

(Crahay, 2019)



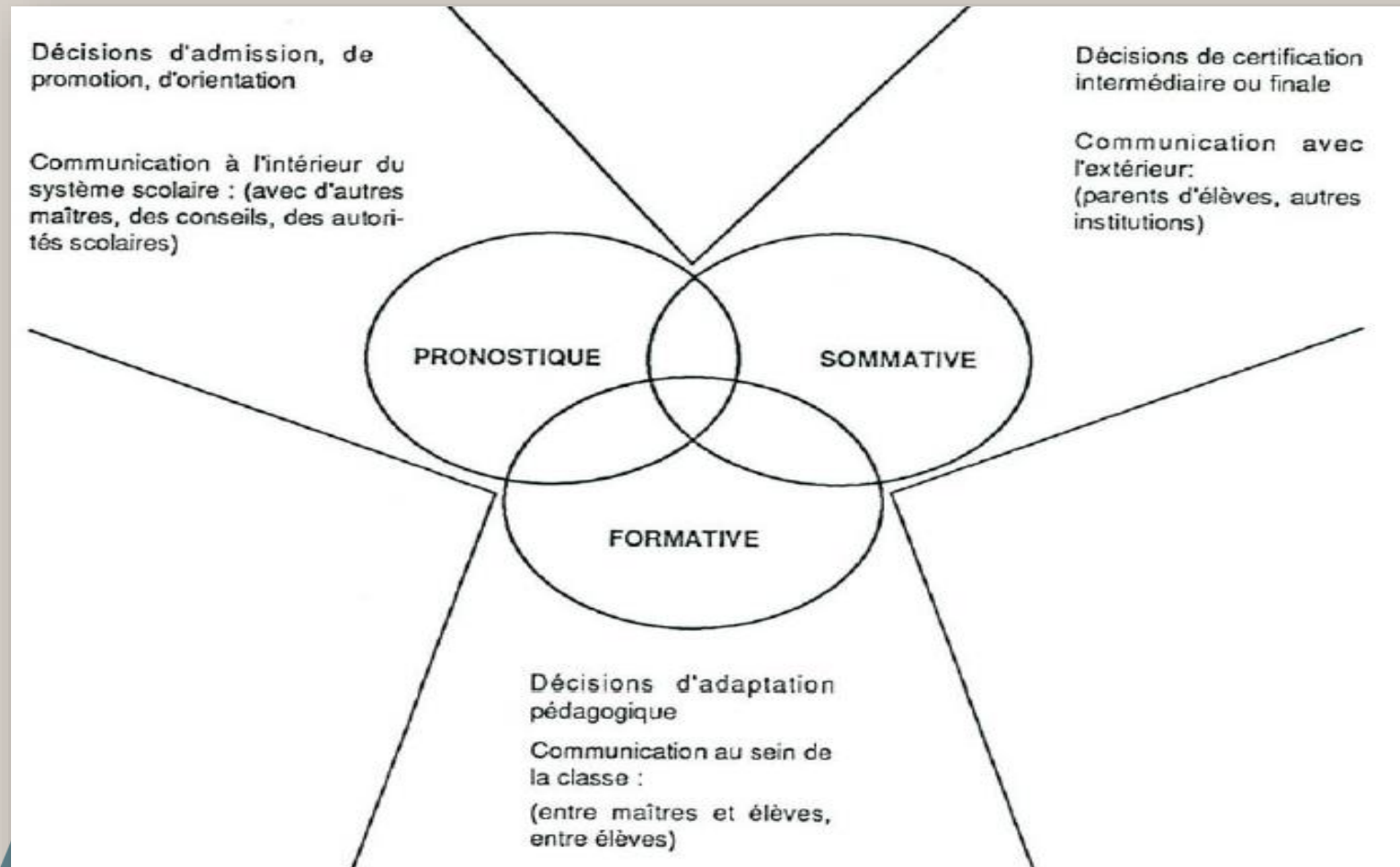
# Des fonctions multiples, mais une visée de régulation

(Crahay, 2019)

Fonction	Définition (Récolte d'informations permettant de ...)	Visée
Formative	Adapter les activités d' <u>enseignement</u> et d' <u>apprentissage</u> afin de <u>favoriser la progression</u> des apprenants vers les objectifs de formation	Régulation (enseignement ou apprentissage)
Sommative (Certificative)	Établir un <u>bilan certifiant</u> les compétences et/ou les connaissances acquises par l'apprenant à la <u>fin d'une étape</u> d'un cycle de formation.	Régulation (atteinte des objectifs)
Pronostique	Guide les <u>décisions d'admission</u> ou d'orientation de l'apprenant vers une nouvelle étape ou un nouveau cursus de formation.	Régulation (parcours de formation)

# Des fonctions multiples, mais une visée de régulation

(Crahay, 2019)

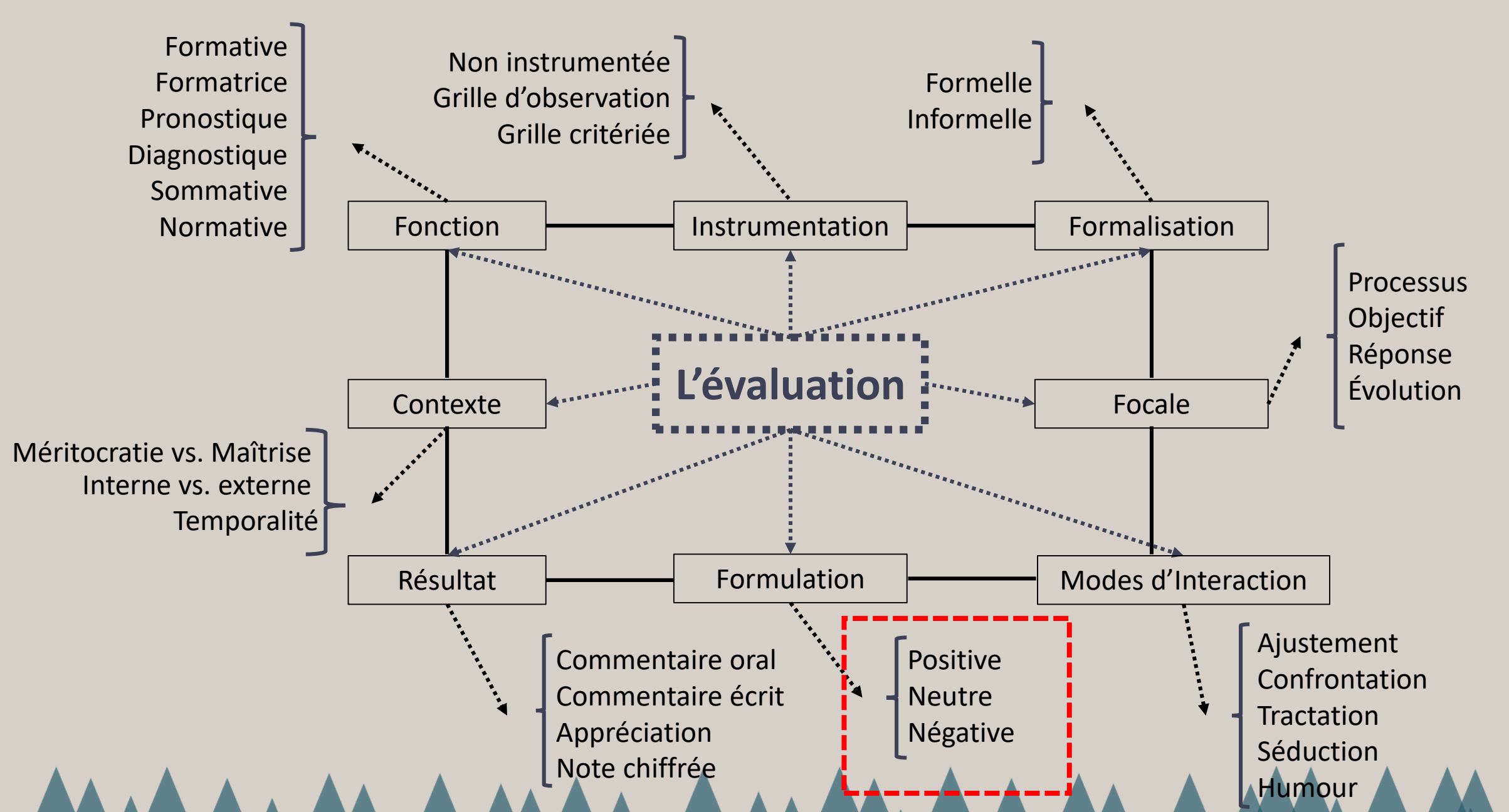




# Des fonctions multiples, mais une visée de régulation

(Crahay, 2019)

Confusion	Exemple
<b>Formative Vs. Pronostique</b>	Réponses à des tests formatifs utilisées pour prédire l'avenir scolaire des élèves
<b>Sommative Vs. Pronostique</b>	Examens pronostics portant sur des compétences issues de programmes antérieurs et non sur les prérequis Épreuves sommatives constitués de questions discriminantes et/ou portant sur des compétences non encore développées
<b>Formative Vs. Sommative</b>	Notation des épreuves formatives Épreuves sommatives calquées sur l'évaluation formative et/ou aboutissant sur de la remédiation



# Le renforcement positif

(Royer et al., 2018)

## Revue systématique

QR : *Est-ce que la pratique du renforcement positif spécifique au comportement peut-être considérée comme fondée sur des preuves ?*

- L'éloge = **subordonnée** au comportement, **spécifique**, **sincère**, **variée**, et **crédible**
- = **facteur malléable** plutôt facteur non contrôlable
- = manière dont le **comportement rencontre l'attente** ou **affecte la réussite**

**Méthode** : recherche dans 21 bases de données bibliographiques → 318 références

recherche à la main → 32 références

- 282 articles exclus pour des raisons de pertinence
- 29 articles exclus après lecture des synthèses
- 33 articles exclus après lecture complète
- 6 articles rencontrent les critères d'inclusion (enseignants ; comportements ; K-12 ; recherche expérimentale)

# Le renforcement positif

(Royer et al., 2018)

## Revue systématique

Étude	Mesure	Estimation
Madsen et al. (1968)	Comportements inappropriés	1.77
Van der Mars (1989)	Comportements hors de la tâche	2.19
Houghton et al. (1990)		1.63
Wheatley et al. (2009)	Comportements inappropriés	1.42
Haydon & Musti-Rao (2011)	Perturbations	3.01
Hollingshead et al. (2016)	Comportements dans la tâche	/

« Les six études examinées montrent que le renforcement positif peut être mis en œuvre avec **un minimum d'effort** et pratiquement **aucune perturbation des enseignements, alors qu'il entraîne des changements rapides dans les résultats des élèves**. D'autres études plus rigoureuses sur le plan méthodologiques sont nécessaires pour déterminer si le renforcement positif comme utilisé dans les écoles traditionnelles de la maternelle à la fin de l'enseignement secondaire supérieur peuvent être considérés comme des pratiques evidence-based. »

# Le renforcement positif

(Fefer et al., 2016)

## Enquête sur les adolescents

QR : *Quelles sont les préférences des élèves adolescents en matière de renforcements positifs et de récompenses relatifs au comportement scolaire ?*

**Méthode** : enquête administrée à 764 élèves de la 5<sup>e</sup> année (P5) à la 12<sup>e</sup> (S6) année d'un district scolaire américain

### Résultats :

- Les adolescents **apprécient** tant les renforcements positifs que les récompenses
- Les adolescents accordent, souvent, **plus d'importance à l'opinion des enseignants** concernant le travail scolaire qu'à l'opinion d'autres (pairs, parents, ...)
- Les **renforcements positifs « silencieux »** (ou privés/discrets) sont plus appréciés que les renforcements positifs publics.
  - Ce dernier constat semble surtout vrai pour les adolescents, moins pour les plus jeunes (Burnett, 2001)

# Le renforcement positif

(Morin, 2022)

- Le renforcement positif **personnel**  
Met l'accent sur les **talents naturels**, les **compétences individuelles**, ...  
« *Tu es vraiment doué en algèbre !* »
- Le renforcement positif fondé sur l'**effort**  
Met l'accent sur le processus d'apprentissage, le temps investi, ... et donc sur des **éléments contrôlables**  
« *Je suis impressionné par le temps que tu as investi pour maîtriser les équations du premier degré.* »
- Le renforcement positif spécifique à un **comportement**  
Met l'accent sur la description du **comportement attendu observé**  
« *Je te félicite d'être resté silencieux pendant que les autres travaillaient.* »

# Le renforcement positif

(Understood for all, 2019)

Caractéristiques	Définition	Contre-exemple
Sincère et véritable	Repose sur une forme de vérité // éviter les exagérations (Brummelman et al., 2014)	
Subordonné au comportement	Dépend intimement du comportement observé	
Spécifique	Décrit exactement le comportement observé et son intérêt	
Pas directif	Ne décrit pas des tâches futures à accomplir	
Centré sur le processus	Renforce la démarche, le progrès, ... plus que le résultat et/ou la comparaison avec les autres	
Immédiat	Est formulé immédiatement après l'observation du comportement	
Cohérent avec les attentes	N'est utilisé que pour valoriser un comportement observé qui correspond au comportement attendu et explicité	
Inattendu	N'est pas prévisible	
Varié	Porte sur l'ensemble des comportements attendus	



# Conclusion

4

4

# Changer l'évaluation pour changer l'école

(Crahay, 2019)

Évaluation dans une perspective ...	
D'excellence	De maîtrise
<p>Évaluation <b>normative</b></p> <ul style="list-style-type: none"><li>→ Classement des individus les uns/aux autres</li><li>→ Distribution gaussienne des notes</li><li>→ Questions discriminantes</li></ul>	<p>Évaluation <b>sommative</b> à référence <b>critériée</b></p> <ul style="list-style-type: none"><li>→ Permettre de situer chaque élève par rapport aux compétences assignées comme objectifs</li></ul>
<p>Évaluation <b>pronostique</b></p> <ul style="list-style-type: none"><li>→ Apte à suivre ?</li><li>→ Sélection.</li></ul>	<p>Évaluation <b>formative</b></p> <ul style="list-style-type: none"><li>→ Comprendre les difficultés éprouvées par l'élève<ul style="list-style-type: none"><li>• Diagnostic</li><li>• Analyse des réponses</li><li>• Interactions avec les élèves</li></ul></li><li>→ Réorienter l'apprentissage et/ou l'enseignement</li></ul>

# Changer l'évaluation pour changer l'école

Conception initiale de Bloom versus conception élargie de l'évaluation formative dans la littérature scientifique francophone (Allal & Mottier Lopez, 2005)

Conception initiale de Bloom	Conception élargie
Insertion de l'EF <u>après</u> la phase d'enseignement	Intégration de l'EF durant <u>tout</u> l'apprentissage
Utilisation de <u>tests</u> formatifs	Utilisation de <u>divers moyens</u> de recueil d'informations
Feed-Back + correction → <u>remédiation</u>	Feed-back + adaptation de l'enseignement → <u>régulation</u>
Gestion de l'EF par l' <u>enseignant</u>	Participation active des <u>élèves</u> à l'EF
Maîtrise des objectifs par <u>tous les élèves</u>	<u>Différenciation</u> de l'enseignement et des objectifs
<u>Remédiation bénéfique</u> aux élèves qui ont été évalués	<u>Régulation à 2 niveaux</u> : pour les élèves évalués et pour les futurs élèves

# Changer l'évaluation pour changer l'école

(Crahay, 2019)

## Évaluation formative formelle

Conçue pour produire des traces et indices sur l'apprentissage de l'élève  
→ *épreuves papier-crayon débouchant sur un feed-back critérié explicite*

### Recueillir

Des informations sur les élèves dans un temps planifié

### Interpréter (analyser)

Les informations recueillies (souvent hors de la présence des élèves)

### Agir

en planifiant une action visant à soutenir l'élève dans l'atteinte des buts d'apprentissage

## Évaluation formative informelle

Traces et indices générées pendant l'activité quotidienne de classe  
→ Pleinement intégrées aux processus d'enseignement et d'apprentissage en contexte de classe

### Solliciter

Des réponses verbales des élèves afin de récolter de l'information

### Reconnaître

Les réponses des élèves au regard des concepts enseignés

### Utiliser

Immédiatement l'information dans le cours des activités continues de la classe

# Bibliographie

5

5

- Allal, L. et Mottier Lopez, L. (2005). Formative assessment of learning : A review of publications in French. In *Formative Assessment - Improving Learning in Secondary Classrooms* (pp. 241-264). Paris, France : OECD-CERI Publication (What works in innovation in education).
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290. <https://doi.org/10.1037/h0055756>
- Autin, F., Batruch, A., & Butera, F. (2019). The function of selection of assesment leads evaluators to artificially create te social class achievement gap. *Journal of Educational Psychology*, 111, 717-735. <https://doi.org/10.1037/edu0000307>
- Boutang, J., & De Lara, M. (2009). *Les biais de l'esprit: comment l'évolution a forgé notre psychologie*. Paris, France : Odile Jacob
- Brummelman, E., Thomaes, S., Orobio de Castro, B., Overbeek, G., & Bushman, B. J. (2014). "That's Not Just Beautiful-That's Incredibly Beautiful!" The Adverse Impact of Inflated Praise on Children With Low Self-Esteem. *Psychological Science*, 25(3). <https://doi.org/10.1177/0956797613514251>
- Burnette, P. C. (2001). Elementary students' preferences for teacher praise. *The Journal of Classroom Interaction*, 36(1), 16-23. Retrieved from: <https://www.jstor.org/stable/23870540>
- Crahay, M. (2019). *Peut-on lutter contre l'échec scolaire ?* (4th ed.). Louvain-la-Neuve, Belgique : de Boeck supérieur.
- De Landsheere, G. (1980). *Examens et évaluation continue. Précis de docimologie*. Bruxelles, Belgique : Labor.
- Durand, T. C. (2019). *Quand est-ce qu'on biaise ?*. Paris, France : HumenSciences (société)
- Durand, T. C. (2021). *La science des balivernes*. Paris, France : HumenSciences (société)

- Dunning, D. (2011). Chapter five - The Dunning-Kruger Effect: on being ignorant of one's own ignorance. *Advances in Experimental Social Psychology*, 44, 247-296. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Fefer, S., DeMagistris, J., & Shuttleton, C. (2016). Assessing adolescent praise and reward preferences for academic behavior. *Translational Issues in Psychological Science*, 2(2); 153-162. <https://doi.org/10.1037/tps0000072>
- Grünbeck, J. (2021, January 2). 72 Biais Cognitifs à Connaître pour mieux Optimiser vos Conversions. Retrieved from: <https://www.convertize.com/fr/71-biais-cognitifs-optimisation-conversions/>
- Kahneman, D. (2012). *Système 1, système 2 : les deux vitesses de la pensée*. Paris, France: Flammarion
- Kahneman, D. Sibony, O., & Sunstein, C. R. (2021). *Noise*. Paris, France : Odile Jacob
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty : heuristics and biases*. Cambridge : Cambridge University Press.
- Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty : heuristics and biases. *Science*, 185, 1124-1131.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6):1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Lafontaine, D., & Monseur, C. (2009). Les évaluations des performances en mathématiques sont-elles influencées par le sexe de l'élève ? *Mesure et Evaluation en Education*, 32(2), 1-106. <https://doi.org/10.7202/1024955ar>
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), 2098-2109. <https://doi.org/10.1037/0022-3514.37.11.2098>



- Morin, A. (2022). *The power of effective praise: a guide for teachers*. Retrieved from: <https://www.understood.org/articles/en/the-power-of-effective-praise-a-guide-for-teachers>
- Nietzsche, F. (1878). *Humain, trop humain*. Paris, France: Société du Mercure de France.
- Rangvid, B. S. (2015). Systematic differences across evaluation schemes and educational choice. *Economics of Education Review*, 48, 41-55. <http://dx.doi.org/10.1016/j.econedurev.2015.05.003>
- Raynal, F., & Rieunier, A. (2014). *Pédagogie, dictionnaire des concepts clés : Apprentissage, formation, psychologie cognitive* (10th ed.). Issy-les-Moulineaux, France : ESF éditeur
- Rot, N., Butas, Z. (1959). Les distributions des notes scolaires comparées aux distributions des résultats obtenus aux tests de connaissances. *Le travail humain*, XXII, 1-2.
- Royer, D. J., Lane, K. L., Dunlap, J. D., & Parks Ennis, R. (2018). A systematic review of teacher-delivered behavior-specific praise on k-12 student performance. *Remedial and Special Education*, 40(2), <http://dx.doi.org/10.1177/0741932517751054>
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54(2), 93-105. <https://doi.org/10.1037/0003-066X.54.2.93>
- Slavin, R. E. (2007). *Education Research in an Age of Accountability*. Boston, MA : Pearson Education.
- Sprietsma, M. (2013). Discrimination in grading: Experimental evidence from primary school teachers. *Empirical Economics*, 45, 523-538. <http://dx.doi.org/10.1007/s00181-012-0609-x>
- Sunstein, C. R., Bobadilla-Suarez, S., Lazzaro, S. C., & Sharot, T. (2017). How People Update Beliefs about Climate Change: Good News and Bad News. *Cornell Law Review*, 102(6), 1431-1443. Retrieved from: <https://scholarship.law.cornell.edu/clr/vol102/iss6/1/>

- Thalmann, Y.-A. (2021). *La tentation du paranormal*. Paris, France : HumenSciences (Débat)
- Understod for All. (2019). *7 ways to praise students*.
- Usabilis. (2018, may 29). *Définition biais cognitive*. Retrieved from: <https://www.usabilis.com/definition-biais-cognitifs/>
- Van Vyve, C. (2021, July 28). Biais de représentativité: gare aux associations trompeuses. Retrieved from: <https://www.easyvest.be/articles/fr/Post/biais-de-representativite-gare-aux-associations-trompeuses>
- Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63-71.