

Lire et interpréter des contributions expérimentales *primaires*

Mattia A. Fritz

27/03/2023

Résumé

Ce document propose des informations utiles pour s'initier à la compréhension des contributions expérimentales *primaires*, c'est-à-dire reportant l'ensemble du processus de génération, récolte, et analyse de données expérimentales. La première partie du document aborde la structure *standard* des contributions expérimentales primaires : Introduction, Méthode, Résultats, et Discussion. La deuxième partie introduit une méthode systématique pour lire, interpréter et éventuellement planifier une recherche expérimentale : la méthode QALMRI, acronyme de Question, Alternatives, Logique, Méthode, Résultats et Inférences.

Introduction

Malgré le fait que les contributions scientifiques adoptant la méthode expérimentale partagent souvent la même structure et des informations similaires, lire, comprendre et interpréter le contenu de ces sources dites *primaires* reste un exercice difficile (Kershaw et al., 2018). Ce document illustre les caractéristiques communes aux contributions expérimentales et propose une méthode pour en faciliter la lecture, l'interprétation et la planification : la méthode QALMRI (Brosowsky et al., 2020), acronyme anglais de *Question, Alternatives, Logic, Method, Results and Inferences*.

Crédits

Ce document reprend et adapte des contenus présents dans d'autres ressources en anglais déjà disponibles sur le sujet, voir par exemple :

- *Learning To Read Scientific Journal Articles* de James Bartlett
- *Research Methods: Lab Manual* de Matthew J. C. Crump

1 Sources expérimentales primaires

On peut diviser les contributions expérimentales en deux catégories :

- Dans une contribution de type *source primaire*, les chercheurs reportent directement l'ensemble du processus expérimental qui a caractérisé l'expérience.
- Dans une contribution de type *source secondaire*, comme les revues systématiques de la littérature, les ouvrages thématiques, ou les méta-analyses, les auteurs utilisent différentes contributions scientifiques pour combiner/comparer les différentes informations.

1.1 Structure d'une source expérimentale primaire

Les contributions expérimentales *primaires* sont assez facilement reconnaissables grâce à une structure qui est très similaire dans les différents domaines/disciplines scientifiques. Cette structure a été adoptée de manière consistante depuis les années 1980 (Sollaci & Pereira, 2004) et est parfois identifiée avec l'acronyme anglais IMRAD : *Introduction, Method(s), Results And Discussion*. Les contributions présentent en général aussi une section *Conclusion* à la fin du texte. De plus en plus souvent, elles proposent également du matériel supplémentaire, incluant par exemple les données récoltées, le code pour traiter ces mêmes données, ou le matériel expérimental adopté (e.g. stimuli, test, ...).

1.2 Introduction

L'introduction sert à justifier la recherche en termes de sa contribution au réseau de connaissances qui composent la littérature scientifique. On peut résumer l'introduction en deux éléments étroitement liées :

1. Qu'est-ce qu'on sait déjà et comment on le sait

Les contribution proposent un aperçu plus ou moins étendu de la littérature existante sur le même axe/sujet de recherche. Si l'introduction prévoit de sous-sections, ces informations se trouvent souvent dans des sous-sections telles que :

- *Theoretical Background* / Cadre théorique : résumé/synthèse des éléments théoriques pertinents dans la recherche, par exemple la définition de concepts mobilisés, des modèles théoriques de références, etc.
- *Related Works* / Contributions de référence : liste de contributions expérimentales qui partagent des éléments sémantiques avec la contribution en objet, souvent selon une approche similarité/contraste. L'objectif est de mettre en évidence des lacunes ou des divergences dans les contributions existantes pour mieux situer

l'apport de l'expérience menée.

2. Qu'est-ce qu'on veut savoir en plus

Les contributions doivent également illustrer leur apport spécifique, ce qu'elles sont censées amener au(x) domaine(s) de recherche. Ces informations se trouvent en général au tout début de l'article, sous la forme d'une problématique/axe de recherche générale, et à la fin de l'introduction. Au début, on définit la *grande question/problématique* de recherche à laquelle plusieurs chercheurs essaient de contribuer et pour laquelle une seule expérience ne peut sûrement pas proposer une réponse à elle toute seule. À la fin de l'introduction, après donc le cadre théorique et les contributions en lien, la contribution spécifie la *question/problématique spécifique* à travers laquelle les chercheurs pensent pouvoir contribuer à la *grande question/problématique*. Souvent, les informations à ce propos se trouvent dans une sous-section nommée :

- *Research Question(s) (and Hypotheses)* / Question(s) de recherche (et hypothèses) : les objectifs spécifiques de la contribution sont définis souvent même en forme interrogative. Si les chercheurs avancent des explications causales potentielles, celles-ci sont formulées en tant qu'hypothèses théoriques.

1.2.1 Exemple de *mini*-introduction

La pensée computationnelle est un terme se référant, de manière plutôt flexible, à un ensemble d'habiletés, compétences et procédures qui facilitent la résolution de problèmes en s'appuyant sur des principes tirés des sciences informatiques (Denning & Tedre, 2019; Fritz & Schneider, 2019; Guzdia, 2008; Wing, 2008). Dans les 20 dernières années, la pensée computationnelle est de plus en plus considérée en tant que sujet à enseigner déjà lors de la scolarisation primaire et secondaire (Bocconi et al., 2016; Grover & Pea, 2013; Resnick, 2017; Sentance et al., 2018). Les activités prévues à cet effet sont souvent divisées en deux catégories : (1) les activités débranchées qui n'utilisent pas des dispositifs numériques mais des modélisations abstraites dérivées d'objets tangibles; et (2) les activités branchées qui, au contraire, s'appuient sur des dispositifs numériques et, souvent, des technologies de l'information et de la communication appliquées à l'enseignement (ibid.). Il n'existe pour l'instant que peu d'études qui comparent directement les bénéfices des deux approches, et ces études s'intéressent principalement aux effets au niveau des apprentissages (Fincher & Robins, 2019; Guzdia, 2015).

Dans cette contribution, l'accent est plutôt porté sur l'intérêt que les types d'activités peuvent susciter à propos d'une éventuelle carrière dans l'un de domaines communément identifiés avec l'acronyme STEM (*Science, Technology, Engineering, and Mathematics*). Nous formulons à cet effet l'hypothèse que des activités branchées peuvent susciter plus d'intérêt, com-

paré aux activités débranchées, en raison de l’inter-médiation d’un dispositif numérique.

1.3 Méthode

La section sur la méthode concerne le micro-monde créé par les chercheurs afin de répondre à la question de recherche spécifique, et tester éventuellement les hypothèses théoriques avancées. La section méthode a deux fonctions principales :

1. Fournir aux lecteurs des informations utiles à évaluer à quel point le micro-monde est suffisamment sensible, fiable et valide pour *produire/représenter* l’effet de l’intervention expérimentale sur le(s) phénomène(s) d’intérêt.
2. Donner tous les détails utiles à répliquer l’expérience par d’autres chercheurs intéressés à refaire exactement la même expérience (dans la perspective de corroborer ou remettre en questions les résultats obtenus) ou des expériences avec quelques variations qui permettent de contribuer d’avantage à la *grande question/problématique*.

Pour atteindre les deux objectifs, la section sur la méthode propose souvent les sous-sections suivantes¹.

1.3.1 Participants

Les chercheurs fournissent des informations à propos des participant-es à l’expérience. Traditionnellement, on indique le nombre total de participant-es, l’âge moyen, le nombre d’hommes/femmes, l’appartenance à une catégorie professionnel/niveau scolaire, ... D’autres informations plus spécifiques peuvent être illustrés, mais les tendances actuelles sont plus inclinées à dévoiler les informations selon une évaluation éthique/scientifique à propos de la pertinence de ces informations (i.e. la possibilité qu’elles influencent les résultats). Dans la sous-section figurent également les modalités de recrutement et l’éventuelle rémunération.

1.3.2 Design expérimental

Les chercheurs indiquent par exemple s’il s’agit d’un plan expérimental ou quasi-expérimental ; intra-sujets, inter-sujets, ou mixte ; à mesure unique ou répétée ; etc. La répartition des participant-es (e.g. aléatoire complète, aléatoire par blocs, ...) est également indiquée. Cette section – ou éventuellement la section sur les participant-es ou sur l’analyse statistique – présente de plus en plus souvent aussi une justification de la taille de l’échantillon. Cette justification peut s’appuyer sur des calculs de puissance statistique pour le type d’analyse à

¹L’ordre peut varier et parfois deux sections sont groupées, par exemple “Participants et design expérimental” ou “Design expérimental et variables”

mener (voir plus bas) ou d’autres critères comme le budget ou le temps à disposition (Lakens, 2022).

1.3.3 Matériel

Les chercheurs illustrent tout le matériel expérimental déployé dans l’expérience, par exemple des instruments de mesure (e.g. eye-tracker, ...) ; des échelles ou questionnaires ; des stimuli ; etc.

1.3.4 Procédure

Le processus de génération/récolte des données de l’expérience est décrit généralement sous forme d’étapes successives, par exemple la préparation du matériel, l’accueil des participantes, les consignes données, les tâches à effectuer, le débriefing final, etc.

1.3.5 Variables

Dans cette sous-section, les chercheurs illustrent clairement :

- la ou les variables indépendantes (VI), c’est-à-dire l’*opérationnalisation* de l’intervention, avec les différentes modalités, ainsi que les différences entre celles-ci.
- toutes les variables mesurées – variables dépendantes (VD), variables contrôlées, ... – et la manière dont elles ont été mesurées (e.g. par rapport au matériel et à la procédure).

1.3.6 Analyse des données (et hypothèses opérationnelles)

Les chercheurs indiquent comment ils vont traiter les données récoltées : quel type de test statistique est prévu, avec quelles critères d’inférences statistiques ; quel logiciel est utilisé ; quelles données vont être considérées inutilisables ; etc.

Si des hypothèses théoriques ont été émises dans l’introduction à complément de la question de recherche spécifique, les chercheurs illustrent les résultats attendus, par exemple en termes de différences entre moyennes (e.g. la modalité A produira un effet plus grand sur la VD par rapport à la modalité B).

1.3.7 Exemple de *mini*-méthode

$N = 364$ élèves du cycle d’orientation ($M_{age} = 13.7$, $SD_{age} = 1.25$) ont suivi individuellement une série d’activités sur la pensée computationnelle pendant un après-midi avec un-e enseignant-e préalablement formé-e aux séquences pédagogiques. La moitié des élèves a été

aléatoirement assignée à des activités débranchées, tandis que l'autre moitié a effectué les activités branchées. Dans les deux cas, les activités ont été conçues et validées par un comité de programme. À la fin de l'activité, les élèves ont rempli une traduction française du *STEM Career Interest Survey* (Kier et al., 2014), un questionnaire avec 44 items qui mesurent l'intérêt pour une carrière dans les domaines STEM (global et avec une sous-échelle pour chaque élément de l'acronyme). Cette contribution s'intéresse principalement à la dimension *technology* et compare donc les deux groupes sur le score de la sous-échelle relative aux technologies. Les scores des élèves ont été analysés en utilisant un *t*-test de Welch unilatéral (Delacre et al., 2017) pour tester l'hypothèse que les participant-es ayant suivi les activités branchées manifestent un intérêt supérieur pour une carrière dans la sphère technologique des STEM comparé aux participant-es ayant suivi les activités débranchées. La taille de l'échantillon a été déterminée sur la base d'un seuil de contrôle de l'erreur de type I $\alpha = 0.05$ (bilatéral) et de type II de $\beta = 0.95$, ainsi qu'une taille minimale de l'effet d'intérêt – *Smallest Effective Size Of Interest* (SESOI) – de Cohen's $d > 0.25$.

1.4 Résultats

La section des résultats reportent de manière assez *brute* les résultats des analyses statistiques descriptives et inférentielles menées sur les données. Il y a en effet l'habitude dans les contributions expérimentales à diviser l'illustration et la discussion des résultats. Ceci a l'avantage de rendre les résultats des tests statistiques identifiables rapidement, par exemple dans une perspective de reproductibilité ou pour les inclure dans une méta-analyse. Il est donc de bonne pratique de reporter tous les indicateurs des tests statistiques (e.g. présence, direction, magnitude et incertitude autour d'un effet) ainsi que des représentations graphiques qui permettent de mieux comprendre les données/résultats.

Dans cette section, les chercheurs peuvent également indiquer des problèmes avec les données ou des critères d'exclusion qui n'ont pas été prévus à l'avance.

1.4.1 Exemple de *mini*-résultats

Les $N = 364$ participant-es ont manifesté un intérêt moyen pour la sous-échelle *technology* du questionnaire STEM-CIS (Kier et al., 2014) de $M = 25.81$ ($SD = 10.22$). Un résultat similaire au score observé par Kier et collègues (*ibid.*) qui a été de $M = 26.27$ ($SD = 10.41$). Dans le groupe branché, les participant-es ont manifesté un intérêt avec $M = 24.24$ ($SD = 9.87$) ; dans le groupe débranché avec $M = 27.37$ ($SD = 10.36$).

Un *t*-test de Welch a été mené avec le score à la sous-échelle *technology* du questionnaire STEM-CIS (*ibid.*) comme variable dépendante et le type d'activité proposé à l'élève comme

variable indépendante (branchée vs. débranchée). Les résultats au test unilatéral indiquent que le groupe activité *branchée* a manifesté un mineur intérêt pour une carrière dans le domaine STEM comparé au groupe *activité débranchée* : $\Delta M = -3.14$, 95% CI $[-5.22, -1.05]$, $t(361.15) = -2.96$, $p = .003$. La différence d'environ 3 points sur la sous-échelle STEM-CIS correspond à une taille de l'effet de Cohen's $d = -0.31$ $[-0.52; -0.10]$.

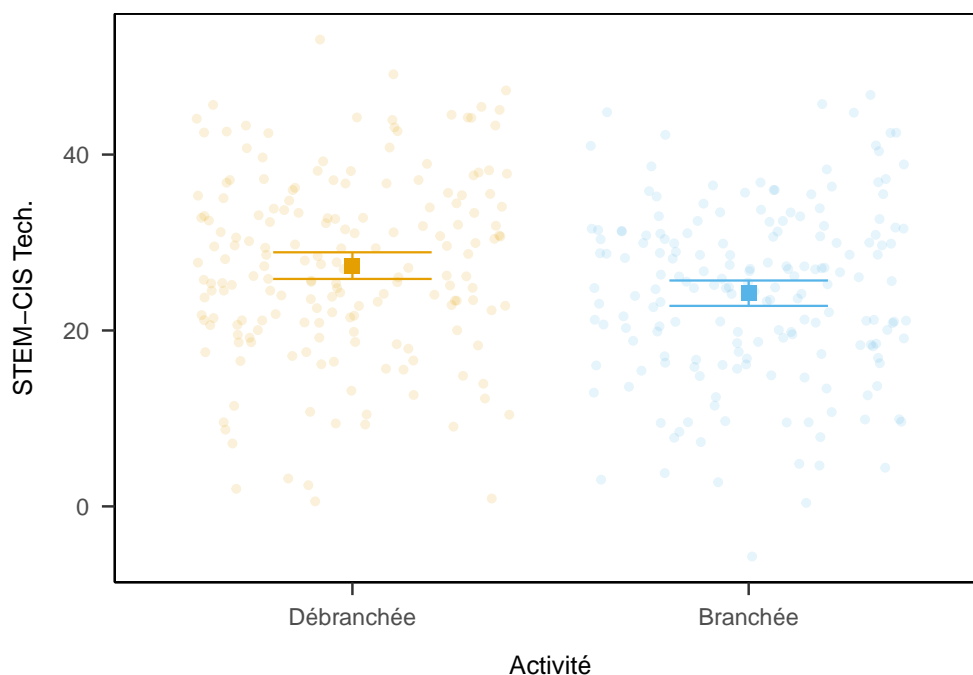


Figure 1: Résultats du questionnaire STEM-CIS pour la sous-échelle technology en fonction du type d'activité branchée (B) ou débranchée (D). Les barres représentent les intervalles de confiance à 95 pourcent.

1.5 Discussion

La discussion intègre les résultats en relation avec la question de recherche et les éventuelles hypothèses émises par les chercheurs. Les chercheurs discutent à quel point les inférences statistiques permettent de répondre à la *question de recherche spécifique*, par exemple en corroborant ou en rejetant les hypothèses. Dans cette section il y a également un retour ou une ouverture sur la littérature scientifique qui essaie de placer les résultats par comparaison/contrastes avec d'autres recherches dans le domaine.

1.5.1 Exemple de *mini-discussion*

Contrairement à notre hypothèse, les élèves ayant suivi des activités débranchées ont manifesté un plus grand intérêt pour une carrière dans le domaine STEM comparé aux élèves

ayant suivi des activités branchées. Notre hypothèse est donc rejetée, avec des résultats qui vont à l'encontre de notre prévision. Cependant, les activités branchées prévues dans notre protocole expérimental ont été perçues comme étant trop difficiles, affectant ainsi très probablement le sentiment d'acceptabilité de la technologie (Scherer et al., 2019; Venkatesh et al., 2003) de la part des élèves. Par reflet, cette difficulté pourrait avoir influencé négativement l'intérêt pour une future carrière dans la sphère technologique des STEM.

1.6 Conclusion

Dans la conclusion, les chercheurs analysent la réponse de l'expérience à la *question de recherche spécifique* en relation avec la *grande question de recherche*. L'apport que l'expérience peut avoir sur l'axe de recherche plus large est évalué de manière critique. Une seule expérience ne produit que des éléments ultérieurs de réflexion, dans l'attente que l'accumulation d'évidence puisse donner des informations plus stables. Les chercheurs indiquent à cet effet des éventuelles limites ou facteurs non considérés qui pourraient expliquer les résultats obtenus, ainsi que des ouvertures pour des recherches ultérieures.

1.6.1 Exemple de mini-conclusion

Cette expérience s'est intéressée à l'effet du type d'activité pédagogique (branchée vs. débranchée) sur l'intérêt de poursuivre une carrière dans le domaine STEM, plus particulièrement dans la sphère des technologies. Les résultats de notre expérience vont à l'encontre de l'hypothèse que les activités branchées pourraient susciter un plus grand intérêt grâce à l'inter-médiation techno-pédagogique d'un outil numérique. Cet effet pourrait néanmoins s'expliquer par un biais du protocole, notamment au niveau de la difficulté des activités branchées. Ces résultats mettent en évidence à nouveau l'importance, et en même temps la difficulté, de trouver des activités branchées adaptées au niveau des élèves du cycle d'orientation (Bocconi et al., 2016; Grover & Pea, 2013). Nous préconisons à cet effet de répliquer l'expérience avec des activités branchées différentes.

2 La méthode QALMRI

La méthode QALMRI (Brosowsky et al., 2020) préconise une manière systématique de penser aux expériences. Cette méthode peut s'appliquer à la lecture et planification des contributions scientifiques primaires surtout de type expérimentale, mais peut s'adapter également aux études basées sur des observations et éventuellement aux méta-analyses.

L'acronyme QALMRI correspond à 6 critères d'analyse d'une contribution expérimentale, c'est-à-dire *Questions, Alternatives, Logic, Method(s), Results and Inferences*. Les mots qui

compose l'acronyme sont très proches au français, ce qui permet de le garder sans la nécessité de convertir des lettres : Questions, Alternatives, Logique, Méthode, Résultats et Inférences.

La liste suivant – librement traduite et adaptée de Brosowsky et collègues (2020) – propose un aperçu des 6 critères, qui seront ensuite abordés de manière plus détaillée. Pour faciliter la lecture, les différentes *entités* sont indiquées aux singulier (question, hypothèse, intervention, phénomène, ...), mais la même contribution expérimentale présente souvent plusieurs *entités*, parfois même plusieurs expériences à la fois.

1. Questions

- Quelle est la question/problématique générale proposée ?
- Quelle est la question/problématique spécifique proposée ?

2. Alternatives

- Quelle est l'hypothèse/explication causale potentielle de l'effet de l'intervention sur le phénomène d'intérêt ?
- Quelle hypothèse/explication causale potentielle pourrait aussi expliquer/prédire l'effet de l'intervention sur le phénomène d'intérêt ?

3. Logique

- Quel est le mécanisme causale à la base de l'hypothèse ?
- Si l'hypothèse est corroborée, qu'est qu'on devrait observer comme résultat ?
- Quel résultat pourrait-on obtenir qui ne corroborerait pas l'hypothèse ?

4. Méthodes

- Comment l'expérience a été mise en place ?
- Quels sont les éléments et les décisions les plus importantes/représentatives ?

5. Résultats

- Comment l'effet de l'intervention a-t-il été évalué/quantifié ?

6. Inférences

- Quelle inférence statistique et *pratique* peut-on tirer des résultats ?
- Sur la base de quel indice est basée l'inférence ?
- Sur la base de quelle logique l'inférence statistique est élargie à l'hypothèse et à la question/problématique de recherche ?

2.1 Questions

La condition naturelle d'un chercheur est le doute, et le doute se traduit par des questions. Est-ce que la technologie peut vraiment faciliter l'apprentissage ? Est-ce qu'une synthèse en forme de carte conceptuelle est plus efficace qu'une synthèse textuelle ? Faut-il enseigner aux élèves la programmation pour améliorer leurs compétences numériques ?

La méthode expérimentale *attaque* les *grandes questions/problématiques de recherche* avec des *questions/problématiques spécifiques* qui se traduisent par la mise en place d'une expérience. L'expérience est censée donner une réponse *spécifique* qui représente un pas en avant dans la direction de la *grande réponse*.

Dans une contribution expérimentale il est important d'identifier les deux niveaux des questions :

- La *grande question/problématique de recherche* est en général liée à un domaine/axe de recherche – uni-, multi- ou inter-disciplinaire – poursuivi par plusieurs chercheurs.
- Si on considère la grande question comme un édifice, alors la *question/problématique spécifique* peut-être considérée comme une brique. En combinant plusieurs questions spécifiques, on essaie d'ériger l'édifice qui sera d'autant plus solide si (1) les briques sont bien intégrées les unes avec les autres ; et (2) des tests sévères ont été menés pour s'assurer que l'édifice ne s'écroule pas – ce qui implique qu'on doit de quelque sorte essayer de le faire écrouler (Mayo, 2018).

Identifier la grande question/problématique de recherche et la question/problématique de recherche spécifique permet de :

- Connaître les intentions de la recherche
- Identifier les concepts principaux mobilisés par la recherche
- Avoir déjà une idée de l'intervention et du phénomène d'intérêt sur lequel un effet est attendu

2.2 Alternatives

Le doute qui amène des chercheurs à concevoir une expérience doit être *plausible*; c'est-à-dire que la réponse ne doit pas être trivial. Par exemple si on s'attend à un effet de l'intervention X_1 plus élevé par rapport à l'intervention X_0 , au moins l'une des alternatives possibles devrait être plausible :

- Les deux interventions produisent des effets plus ou moins similaire, ou dont la diffé-

rence peut être considérée négligeable d'un point de vue *pratique*, c'est-à-dire $X_0 \approx X_1$

- L'intervention X_0 produit un effet plus élevé par rapport à l'intervention X_1

La plausibilité des alternatives est notamment fondamentale par exemple lorsque l'expérience est orientée vers le test d'une théorie ou de deux théories *en compétition*.

Dans les alternatives on peut également considérer d'autres facteurs potentiels – différents de l'intervention – qui pourraient expliquer l'effet sur le phénomène d'intérêt (e.g. variables cachées, médiateurs, ...).

2.3 Logique

La partie logique d'une contribution expérimentale concerne :

- L'identification claire et précise des alternatives/différences dans les interventions X_0 , X_1 , X_2 , ... X_j
- La raison, en termes d'explication causale potentielle, pour laquelle les interventions produisent des effets différents, des effets similaires, ou pas d'effet du tout.

L'analyse logique d'une contribution expérimentale peut se faire à l'aide d'opérateurs logiques utilisés également dans la programmation :

Si p alors q

Par exemple :

- ***Si*** les cartes conceptuelles renforcent davantage les liens sémantiques entre concepts comparé à une synthèse écrite, ***alors*** les participant-es dans la condition X_{carte} auront des meilleures performances au test de compréhension par rapport aux participant-es dans la condition $X_{synthese}$
- ***Si*** des vidéos de plus courte durée permettent d'éviter des phénomènes de distraction chez les apprenant-es, ***alors*** les participant-es ayant observé trois vidéos de 5 minutes (X_{3x5min}) auront des meilleures performances au test de rétention mnésique par rapport aux participant-es ayant observé le même contenu dans une vidéo de 15 minutes ($X_{1x15min}$).

2.4 Méthodes

Le critère relatif à la méthode correspond à une compréhension/description du micro-monde dans lequel :

- L'intervention X devient la variable indépendante (VI) avec les différents facteurs/modalités
- Le phénomène d'intérêt Y devient la variable dépendante (VD) avec son mécanisme de mesure sensible, fiable et valide

Les informations saillantes qui caractérisent le micro-monde sont également à noter, comme par exemple le public cible/participant-es, le type de tâche/stimuli, le matériel adopté, etc.

2.5 Résultats

Les résultats des expériences sont généralement divisés en deux types :

- La description de l'échantillon, c'est-à-dire les données observées/mesurées dans l'expérience
- Les statistiques inférentielles qui *généralisent* le micro-monde au macro-monde

Dans les deux cas, l'analyse de résultats essaie d'établir :

- La présence de patterns dans les données au niveau de l'échantillon général et des stratifications/groupements par facteurs/conditions
- La plausibilité/fiabilité des résultats observés (e.g. *too good to be true*)
- La pertinence des tests/indices statistiques
- L'utilité des représentations graphiques de support (tableaux, graphiques, ...)

2.6 Inférences

Les inférences concerne principalement la capacité de l'expérience à répondre à la question/problématique spécifique. En deuxième lieu, l'inférence s'applique à la généralisation/intégration de l'expérience avec la grande question/problématique de recherche.

Des éléments corollaires à tenir en compte concernent des éventuelles problèmes survenus pendant l'expérience, la possibilité qu'il existe d'autres facteurs non mesurés qui pourraient influencer/expliciter/annuler l'effet de l'intervention, ainsi que des modifications à apporter pour des futures expériences sur le même thème.

3 Conclusion

Ce document a fourni des informations utiles pour s'initier à la lecture et interprétation de contributions expérimentales primaires. D'abord, la structure *classique* des contributions

a été illustrées en fonction des objectifs et des pratiques courantes dans l'exposition des contenus pour chaque section. Ensuite, la méthode QALMRI a été proposée en tant que pratique systématique pour lire et interpréter (et éventuellement planifier) des contributions expérimentales.

Références

- Bocconi, S., Chiocciariello, A., Dettori, G., Ferrari, A., Engelhardt, K., Kampylis, P., & Punie, Y. (2016). Developing Computational Thinking : Approaches and Orientations in K-12 Education. *Proceedings EdMedia 2016, June*, 1-7. <https://doi.org/10.2791/792158>
- Brosowsky, N., Parshina, O., Locicero, A., & Crump, M. J. C. (2020). *Teaching undergraduate students to read empirical articles: An evaluation and revision of the QALMRI method*. <https://doi.org/10.31234/osf.io/p39sc>
- Delacre, M., Lakens, D., & Leys, C. (2017). Why Psychologists Should by Default Use Welch's *t*-test Instead of Student's *t*-test. *International Review of Social Psychology*, 30(1), 92. <https://doi.org/10.5334/irsp.82>
- Denning, P. J., & Tedre, M. (2019). *Computational Thinking*.
- Fincher, S. A., & Robins, A. V. (Éds.). (2019). *The Cambridge Handbook of Computing Education Research*. Cambridge University Press.
- Fritz, M. A., & Schneider, D. K. (2019). *Pensée computationnelle avec JavaScript : le cours STIC I*.
- Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, 42(1), 38-43. <https://doi.org/10.3102/0013189X12463051>
- Guzdial, M. (2008). Paving the way for computational thinking. *Communications of the ACM*, 51(8), 25. <https://doi.org/10.1145/1378704.1378713>
- Guzdial, M. (2015). *Learner-Centered Design of Computing Education. Research on Computing for Everyone*. Morgan & Claypool.
- Kershaw, T. C., Lippman, J. P., & Fugate, J. M. B. (2018). Practice makes proficient: teaching undergraduate students to understand published research. *Instructional Science*, 46(6), 921-946. <https://www.jstor.org/stable/45213864>
- Kier, M. W., Blanchard, M. R., Osborne, J. W., & Albert, J. L. (2014). The Development of the STEM Career Interest Survey (STEM-CIS). *Research in Science Education*, 44(3), 461-481. <https://doi.org/10.1007/s11165-013-9389-3>
- Lakens, D. (2022). Sample Size Justification. *Collabra: Psychology*, 8(1), 33267. <https://doi.org/10.1525/collabra.33267>
- Mayo, D. G. (2018). *Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars* (p. i-i). Cambridge University Press.

- Resnick, M. (2017). *Lifelong Kindergarten. Cultivating Creativity through Projects, Passion, Peers, and Play*. MIT press.
- Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Sentance, S., Barendsen, E., & Schulte, C. (Éds.). (2018). *Computer science education: perspectives on teaching and learning in school*. Bloomsbury Academic.
- Sollaci, L. B., & Pereira, M. G. (2004). The introduction, methods, results, and discussion (IMRAD) structure: a fifty-year survey. *Journal of the Medical Library Association*, 92(3), 364-371. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC442179/>
- Venkatesh, Morris, Davis, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425. <https://doi.org/10.2307/30036540>
- Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 366(1881), 3717-3725. <https://doi.org/10.1098/rsta.2008.0118>