



# The Development of In-Person and Remote Literacy Assessments in Filipino and English

Kathrina Lorraine Lucasan<sup>1,2</sup>, Dina Ocampo<sup>2</sup>, and Margaret Mary Rosary Carmel Fua<sup>2</sup>

<sup>1</sup> Ateneo de Manila University

<sup>2</sup> University of the Philippines Center for Integrative and Development Studies-  
Research Program (UP CIDS-ERP)

Education



# Introduction

- Assessment is “the process of gathering, describing or quantifying information about learner performance” (Rovai, 2000, p. 142).
- Gonzales (2023) describes how assessment holds various meanings
  - Used by teachers to **grade student assignments** (Harlen, 2010; Mertler, 2003; Schuh & Upcraft, 2001; Unal & Unal, 2019)
  - **Standardize testing** imposed in schools (Morgan, 2016; Paul, 2015; Salinas & Guajardo, 2022; Stiggins & Chappuis, 2005)
  - Any activity designed to collect information to be used as **feedback to modify teaching and learning activities** (Black & William, 2003; Chappuis, 2009, 2004; Hattie & Timperley, 2007)
  - **Improve instruction and students’ performance** (Arnaiz-Sanchez et al., 2020; Cohen & Hill, 2000; Dunn et al., 2017; Ramberg et al., 2019)



# Introduction

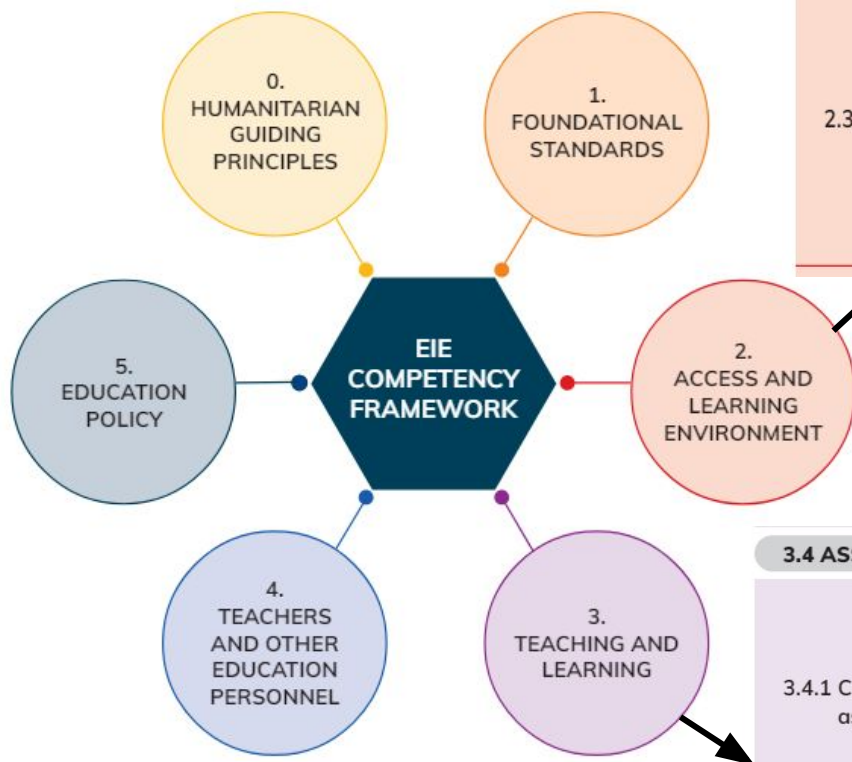
- The pandemic has changed the way that education is implemented (Gonzales, 2023), including the administration of assessments (Fagell, 2020; Gonzales, 2023; Hurtz & Weiner, 2022; Krach, et al., 2020).
- The general principles surrounding assessment of student performance **do not change** in an online setting.
- Assessments in an online setting **become even more important** when there are no face-to-face interactions between teachers and students (Rovai, 2000).

# Introduction

- The Inter-agency Network for Education in Emergencies (INEE) has a set of “Minimum Standards.”
- These standards aim to “enhance the quality of educational preparedness, response and recovery, increase access to safe and relevant learning opportunities, and ensure accountability in providing these services in crisis contexts” (INEE, 2020, p. 5).
- Each of the domains have corresponding “competencies” relevant to the domain.



# Introduction



## 2.3 FACILITIES AND SERVICES

2.3.1 Learning environment	Facilities and services are <b>safe</b> , inclusive and accessible, and meet minimum national and/or Sphere standards (Sphere Association, 2018)	Describe the characteristics of safe and inclusive learning environments that meet minimum national and/or Sphere standards	Work with community and children to identify and reduce risks and create safe, inclusive learning environments that meets minimum national and/or Sphere standards	Design and train others on risk reduction and participatory school improvement methodology which promote and meets minimum national and/or Sphere standards for safe and inclusive learning environments
----------------------------	--	---	--	--

## 3.4 ASSESSMENT OF LEARNING OUTCOMES

3.4.1 Classroom based assessment	Teachers' capacity to use both formative and summative assessment as appropriate to the context	Explain the purpose of and difference between formative and summative assessment. Identify <b>different methods of administering assessments</b>	Build teacher capacity to use a range of formative and summative assessment strategies and to record learner progress. Facilitate learner access to appropriate placement and progress tests and any relevant exams	Collaborate with other stakeholders in development and harmonisation of placement tests, promotion tests and level exams. Advocate for and develop tools to support greater use of formative assessment
3.4.2 Measuring learning	Standard proficiency assessments used to identify learning outcomes	Identify standard tools and processes for assessing learning outcomes. Explain when and why different tools should or should not be used	Select, <b>adapt and use appropriate and contextualised tools</b> and processes to assess learning	Use the results from assessments to design instructional programs and assess progress. Train others on how to develop, adapt and use culturally relevant and valid tools

# Introduction

- According to Bearman et al.'s (2016) Assessment Design Decisions Framework, six key aspects should be considered when developing or modifying assessments.





# Introduction

- Test scores from in-person and online assessments have been found to be equivalent (Dan & Pham, 2023; Hall et al., 2001; Hamner et al., 2021; Hodge et al., 2019; Krach et al., 2020; Magimairaj et al., 2022; Wright, 2020).

# Introduction

- The University of the Philippines Center for Integrative and Development Studies-Education Research Program (UP CIDS-ERP) has developed **a battery of literacy assessments for Kindergarten to Grade 3 (K to 3), originally developed for in-person administration** (UP CIDS-ERP, 2023). This was the Multi-Literacy Assessments for Filipino Children (MLAF).







# Introduction

Experiences from the pandemic have highlighted the need for **alternative tools for assessment** in case of long school disruptions similar to the pandemic. This facilitated the development of MLAF 2.0, which not only has a version for in-person administration, but also for remote modes of administration.

1. Alphabet Knowledge\*
2. Book and Print Knowledge\*
3. Figure Copying\*
4. Fluency and Reading Comprehension
5. Grammar
6. Handwriting\*
7. Listening Comprehension
8. Logical Sequencing\*
9. Phonics and Word Recognition
10. Phonics and Non-Word Recognition
11. Phonological Awareness
12. Rapid Automatized Naming
13. Spelling
14. Vocabulary

\*administered only in the language preferred by the child

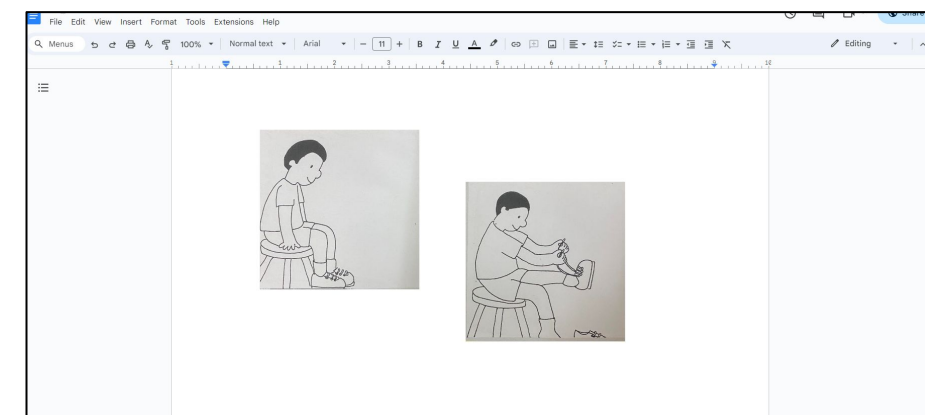
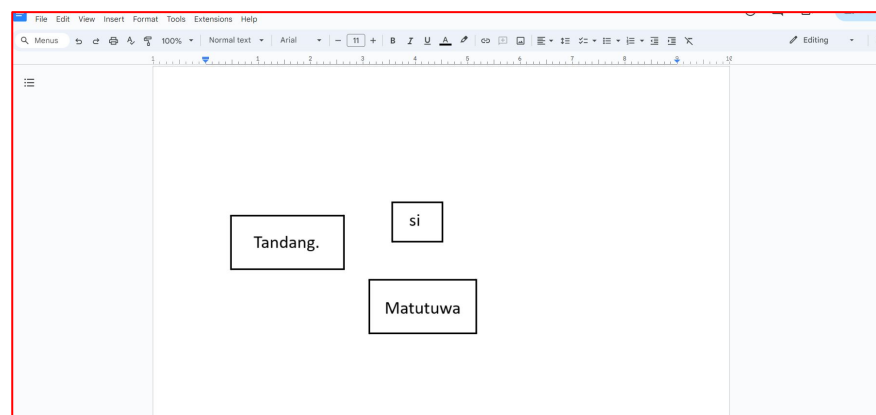
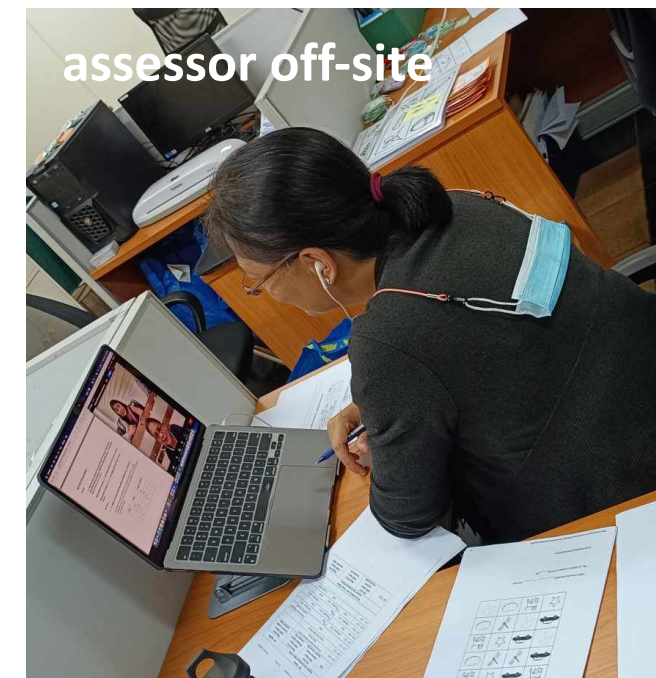
# MLAF (Face-to-Face) In-Person Mode

- Administered to learners by an assessor (both on-site)
- Assessor marks the score sheet to capture the responses of the child
- Some components may be administered to a group of students (Vocabulary, Listening Comprehension, Spelling, & Handwriting)



# MLAF (Visual Remote) Online Mode

- Takes place via video conference using Zoom or FB Messenger
- Use of computer or smartphone
- Use of Google Doc or other file sharing platforms for components involving real-time manipulation of word or picture cards
- On-site assistance



# Alphabet Knowledge (in-person and online)



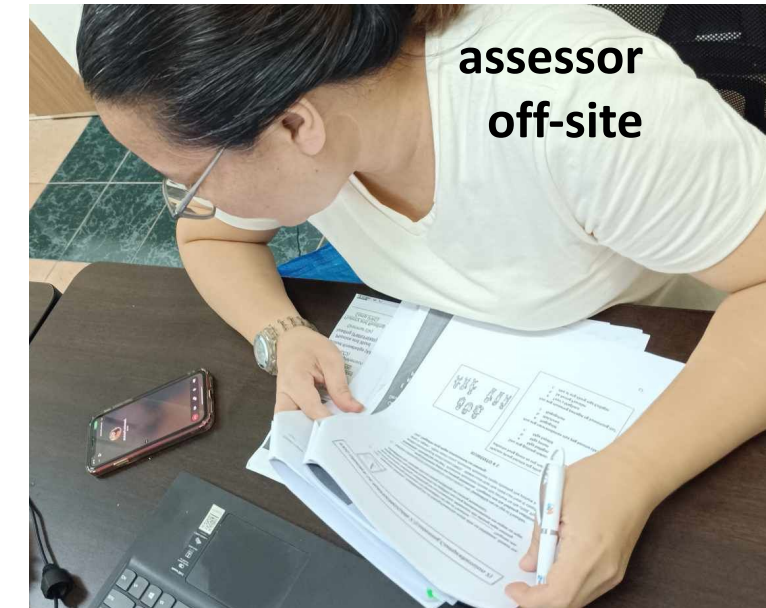
**in-person**



**online**

# MLAF (Telephony)

- Takes place via audio call (call is placed on speaker phone)
- Use of manipulatives introduced & demonstrated by assistant on-site





# Comparison: Roles & Distribution of Tasks

	Child assent form	Intro/ Rapport-building	Giving of prompt	Demo	Presenting stimulus	Distribution/ Collection of answer sheet	Audio recording	Recording of time	Scoring
In-person	Assessor								
Online	Assistant	Assessor	Assessor	Assessor/ Assistant	Assessor	Assistant	Assessor	Assistant	Assessor/ Assistant
Tele-phony	Assistant	Assessor	Assessor	Assistant	Assistant	Assistant	Assistant	Assistant	Assistant/ Assessor



# Introduction

- The study extends the literature on the equivalence of in-person and online literacy assessments in the Philippine context.
- It will inform the further development and refinement of literacy assessments for Filipino children for in-person and online modalities.



# Research Design

- Quantitative, correlational study (Creswell, 2012)
- Focused on determining the relationships of the in-person and online version assessments administered





# Hypothesis

- There is a significant difference between literacy assessment results administered in-person and online for the following:
  1. Alphabet Knowledge
  2. Book and Print Knowledge
  3. Figure Copying
  4. Fluency and Reading Comprehension
  5. Grammar
  6. Handwriting
  7. Listening Comprehension
  8. Logical Sequencing
  9. Phonics and Word Recognition
  10. Phonics and Non-Word Recognition
  11. Phonological Awareness
  12. Rapid Automated Naming
  13. Spelling
  14. Vocabulary



# Assessment Data

- Beginning-of-year MLAF K to 3 assessment data (in-person and online)
  1. Alphabet Knowledge (64)
  2. Book and Print Knowledge (13)
  3. Figure Copying (8)
  4. Fluency and Reading Comprehension (20)\*
  5. Grammar (6)\*
  6. Handwriting (32)
  7. Listening Comprehension (9)\*
  8. Logical Sequencing (6)
  9. Phonics and Word Recognition (26)\*
  10. Phonics and Non-Word Recognition (11)\*
  11. Phonological Awareness (28)\*
  12. Rapid Automatized Naming (48)\*
  13. Spelling (26)\*
  14. Vocabulary (9)\*

\*total number of items PER language

Total of 23 Assessments (Filipino Assessments: 9; English Assessments: 9; Non-language Assessments: 5)

Overall total of 449 items (Fil: 183; Eng: 183; NL: 123)

# Participants

## ➤ Number of Learners (by grade level and by hub)

	Hub 1		Hub 2		Hub 3		Hub 4		Total	
	F2F	VR	F2F	VR	F2F	VR	F2F	VR	F2F	VR
Kindergarten	1	0	1	2	1	0	0	0	3	2
Grade 1	4	2	1	1	1	2	0	0	6	5
Grade 2	1	0	3	2	1	1	2	0	7	3
Grade 3	1	3	0	1	2	2	1	3	4	9
Total	7	5	5	6	5	5	3	3	20	19

- *Number of learners per modality: In-person=20; Online=19*



# Method

- The t-Test was used to determine the equivalence of tests (Still & Still, 2015).
  - t-test/Mann-Whitney Test/Welch's t-Test (Skaik, 1969)

# Findings



## Scale Reliability Statistics

---

	Cronbach's $\alpha$
Filipino Assessments (183 items)	.909
English Assessments (183 items)	.911
Non-Language Assessments (123 items)	.847

# Method



The normality and homogeneity of variances assumptions were tested.

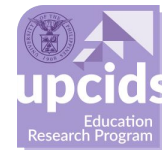
Met Both Assumptions (5)		Met Homogeneity of Variances Assumption Only (17)		Met Normality Assumption Only (2)	
Conduct Student's t-Test		Conduct Mann-Whitney Test		Conduct Welch's t-Test	
Listening Comprehension (F)	Listening Comprehension (E)				
Phonological Awareness (F)					Phonological Awareness (E)
Vocabulary (F)	Vocabulary (E)				
				Book & Print Knowledge	
		Alphabet Naming			
		Alphabet Sounding			
		Figure Copying			
		Grammar (F)	Grammar (E)		
		Fluency & Reading Comprehension (F)	Fluency & Reading Comprehension (E)		
		Handwriting			
		Logical Sequencing			
		Phonics & Word Recognition (F)	Phonics & Word Recognition (E)		
		Phonics & Non-Word Recognition (F)	Phonics & Non-Word Recognition (E)		
		Rapid Automated Naming (F)	Rapid Automated Naming (E)		
		Spelling (F)	Spelling (E)		

# Findings

Student's t-Test		Mann-Whitney Test		Welch's Test	
Assessment	p value	Assessment	p value	Assessment	p value
LC (Fil)	.477				
PA (Fil)	.595			PA (Eng)	.967
LC (Eng)	.003				
V (Fil)	.410	PWR (Fil)	.943		
V (Eng)	.326	PWR (Eng)	.930		
		PNWR (Fil)	.926		
		PNWR (Eng)	.639		
		S (Fil)	.834		
		S (Eng)	.748		
		G (Fil)	.593		
		G (Eng)	.193		
		RAN (Fil)	.059		
		RAN (Eng)	.888		
		FRC (Fil)	.540		
		FRC (Eng)	.710		
		AK N	.899		
		AK S	.383		
		HW	.966		
		FC	.778		
		LS	.324	BPK	.410



ATENEO DE MANILA  
UNIVERSITY



# Findings

## Group Descriptives

	Group	N	Mean	Median	SD	SE
LC (Fil)	F2f	20	64.444	72.222	28.293	6.327
	Vr	19	58.480	55.556	23.083	5.296
PA (Fil)	F2f	20	49.286	48.214	25.003	5.591
	Vr	19	45.301	39.286	21.097	4.840
LC (Eng)	F2f	20	<b>60.556</b>	61.111	17.834	3.988
	Vr	19	<b>39.766</b>	33.333	22.625	5.190
V (Eng)	F2f	20	60.000	61.111	20.200	4.517
	Vr	19	53.801	55.556	25.995	5.964
V (Fil)	F2f	20	73.889	77.778	12.105	2.707
	Vr	19	77.778	77.778	12.284	2.818





# Discussion

- The t-Test results showed **no significant differences between the in-person and online modalities**, except for Listening Comprehension in English; This may be due to technical difficulties experienced during the online assessments.
- The findings **confirm** the studies of Dan and Pham (2023), Hall et al. (2001), Hamner et al. (2021), Hodge et al. (2019), Krach et al. (2020), Magimairaj et al. (2022), and Wright (2020).
- The online version of the MLAF could be used when in-person assessments are not possible, and may be used to satisfy the aspects of safety and contextualization of the INEE (2000) and Bearman et al.'s (2016) frameworks.



# Recommendations

- Revisit/Review the following MLAF administration aspects:
  - quality of the recorded story when shared via Zoom;
  - general visual remote pre-administration and during administration protocols;
  - specific protocols for playing of audio recording of story during the listening comprehension test (ensure maximum volume setting, no/limited background noise)
- Explore having learners “point” to/indicate/circle answers onscreen, instead of answering sheets offline.
- Administer assessments to more children to be able to employ other statistical treatments.
- Include the telephony mode in the analysis.



# References

- Alexander, M. W., Bartlett, J. E., Truell, A. D., & Ouwenga, K. (2001). Testing in a Computer Technology Course: An investigation of equivalency in performance between online and paper and pencil methods. *Journal of Career and Technical Education*, 18(1). <https://doi.org/10.21061/jcte.v18i1.600>
- Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016). Support for assessment practice: developing the Assessment Design Decisions Framework. *Teaching in Higher Education*, 21(5), 545–556. <https://doi.org/10.1080/13562517.2016.1160217>
- Creswell, J. W. (2012) *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (Fourth Edition). Boston: Pearson.
- Dam, O., & Pham, G. (2023). Remote First-Language Assessment: Feasibility study with Vietnamese bilingual children and their caregivers. *Language, Speech & Hearing Services in Schools* (Print), 54(2), 618–635. <https://doi.org/10.1044/2023.lshss-22-00123>
- de Winter, J.C.F. (2019) "Using the Student's t-test with extremely small sample sizes," *Practical Assessment, Research, and Evaluation*, 18(10). <https://doi.org/10.7275/e4r6-dj05>
- Fagell, P. L. (2020). Career Confidential: Teacher wonders how to help students during coronavirus shutdown. *Phi Delta Kappan*, 101(8), 67–68. <https://doi.org/10.1177/0031721720923799>
- Gonzales, R. D. (2023). From Face-to-Face to Virtual Student Assessment: Changes in Student Assessment Practices during COVID-19 Among Filipino Teachers. *European Journal of Education and Pedagogy*, 4(1), 159–164. <https://doi.org/10.24018/ejedu.2023.4.1.579>
- Hamner, T., Salorio, C. F., Kalb, L. G., & Jacobson, L. A. (2021). Equivalency of In-Person versus Remote Assessment: WISC-V and KTEA-3 performance in clinically referred children and adolescents. *Journal of the International Neuropsychological Society* (Print), 28(8), 835–844. <https://doi.org/10.1017/s1355617721001053>
- Hodge, M.A., Sutherland, R., Jeng K, et al. (2019). Agreement between telehealth and face-to-face assessment of intellectual ability in children with specific learning disorder. *Journal of Telemedicine and Telecare*, 25 (7), 431-437. <https://doi.org/10.1177/1357633X18776095>
- Inter-agency Network for Education in Emergencies. (2020). *Education in Emergencies Competency Framework*. New York: New York. [https://inee.org/sites/default/files/resources/INEE\\_Compentency\\_Framework\\_English\\_2D%20Studio%20v3.1%20LowRes.pdf](https://inee.org/sites/default/files/resources/INEE_Compentency_Framework_English_2D%20Studio%20v3.1%20LowRes.pdf)
- Krach, S. K., Paskiewicz, T. L., & Monk, M. M. (2020). Testing Our Children When the World Shuts Down: Analyzing Recommendations for Adapted Tele-Assessment during COVID-19. *Journal of Psychoeducational Assessment*, 38(8), 923-941. <https://doi.org/10.1177/0734282920962839>
- Magimairaj, B. M., Vaughn, S., Gillam, S. L., Vaughn, S., Roberts, G., Fall, A., & Gillam, R. B. (2022). Online Administration of the Test of Narrative Language–Second Edition: Psychometrics and Considerations for Remote Assessment. *Language, Speech & Hearing Services in Schools* (Print), 53(2), 404–416. <https://doi.org/10.1044/2021.lshss-21-00129>
- Rovai, A. P. (2000). Online and traditional assessments: what is the difference? *The Internet and Higher Education*, 3(3), 141–151. [https://doi.org/10.1016/s1096-7516\(01\)00028-8](https://doi.org/10.1016/s1096-7516(01)00028-8)
- Skaik, Y. (1969). The bread and butter of statistical analysis “t-test”: Uses and misuses. *Pakistan Journal of Medical Sciences*, 31(6). <https://doi.org/10.12669/pjms.316.8984>
- Still, M. L., & Still, J. D. (2015). Contrasting Traditional In-Class Exams with Frequent Online Testing. *Journal of Teaching and Learning With Technology*, 4(2), 30–40. <https://doi.org/10.14434/jotlt.v4n2.13481>
- University of the Philippines Center for Integrative and Development Studies-Education Research Program. (2023.) “Proceedings 2023. A Roundtable Discussion on Literacy Assessments for Grades 4 to 12. Emerging Themes and Framework for 21st Century Learning.” Quezon City: University of the Philippines Center for Integrative and Development Studies.
- Wright, A. J. (2020). Equivalence of remote, digital administration and traditional, in-person administration of the Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V). *Psychological Assessment*, 32(9), 809–817. <https://doi.org/10.1037/pas0000939>



# Thank you for listening!

[kmlucasan@up.edu.ph](mailto:kmlucasan@up.edu.ph)

[erp.cids@up.edu.ph](mailto:erp.cids@up.edu.ph)