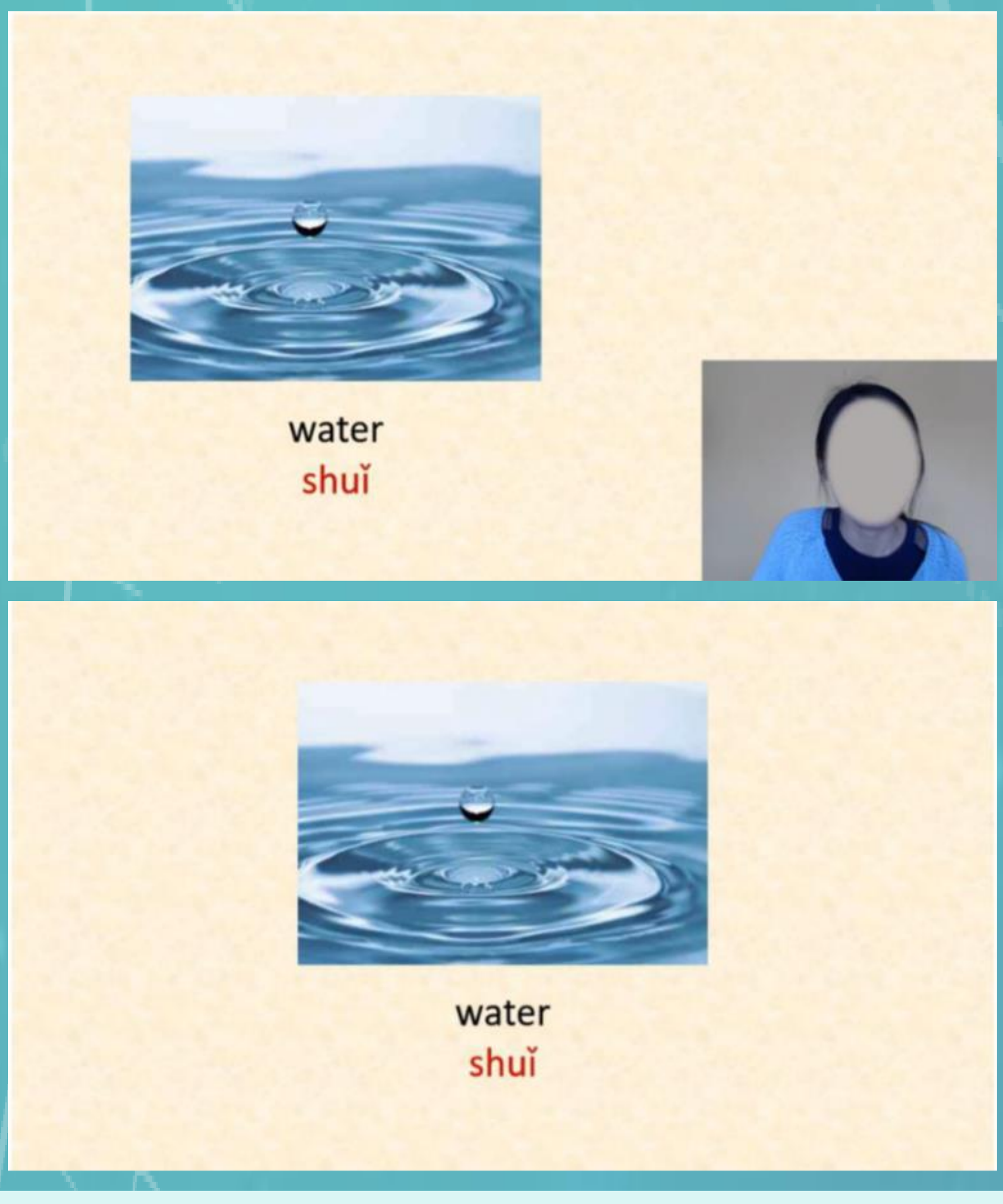
A Systematic Review of Eye Tracking Studies on the Impact of Instructor Presence in Educational Videos

Yen Ying Ng; Centre for Language Evolution Studies, Nicolaus Copernicus University in Torun, Poland; nyysang@hotmail.com
Noah L. Schroeder; Department of Computer and Information Science and Engineering, University of Florida, USA; schroedern@ufl.edu

1 Introduction

Social cues from human instructors can activate social schemata in long-term memory, improving cognitive processes, motivation, and emotional engagement, which in turn enhances learning outcomes in digital environments (Schneider et al., 2021). Therefore, including a human instructor in instructional videos is seen as beneficial (Beege et al., 2023). However, the effectiveness of this design feature remains the subject of ongoing research.

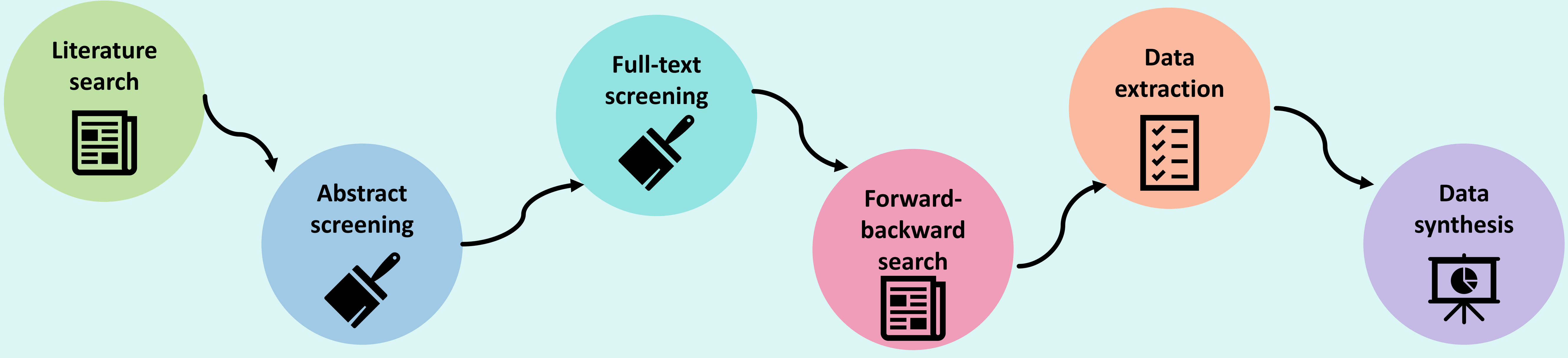
[1] S. Schneider, M. Beege, S. Nebel, L. Schnaubert, & G. D. Rey, "The cognitive-affective-social theory of learning in digital environments (CASTLE)," Educational Psychology Review, 2022, Vol. 34, pp. 1–38. https://doi.org/10.1007/s10648-021-09626-5
[2] M. Beege, N. L. Schroeder, S. Heidig, G. D. Rey, S. Schneider, "The instructor presence effect and its moderators in instructional video: A series of meta-analyses," Educational Research Review, 2023, Vol. 41, 100564. https://doi.org/10.1016/j.edurev.2023.100564



2 Research Questions

(1) How does instructor presence affect learning outcomes?
(2) How are eye-tracking data interpreted in these studies?
(3) How does an instructor influence learners' gaze behavior?
(4) What is the relationship between learners' gaze behavior and learning outcomes?
(5) What instructional design guidelines can be derived from the reviewed studies?

3 Methods



4 Preliminary Results

Study	Study Design	Independent Variables	Learning Outcomes	Eye-tracking Results (Visual Attention)
Pi & Hong (2016)	Between-subjects; 4 groups 1. Instructor-absent (n = 24) 2. Instructor-present (n = 24) 3. Instructor-present without PPT (n = 22) 4. Classroom (n = 24)	Instructor presence*	*Combine retention and transfer* IP > IA ≈ classroom > IP without PPT	*Fixation count* Instructor > content *Mean fixation duration* Content > instructor *Total dwell time* Instructor > content Note: On average, learners spent 62.3% of their time focusing on the instructor.
Wang & Antonenko (2017)	Not clearly described Instructor presence (present vs absent) Content difficulty (easy vs difficult)	Instructor presence Content difficulty	*Retention* Easy + IP > easy + IA Difficult + IP ≈ difficult + IA *Transfer* Easy/difficult + IP ≈ easy/difficult + IA	*Fixation count, total dwell time, total dwell time %* Easy ≈ difficult *Fixation count %* Easy > difficult Note: Learners in the easy topic condition spent an average of 26% of their time looking at the instructor, while those in the difficult topic condition spent 22%.
Zhang et al. (2021)	Between-subjects; 4 groups 1. Instructor-absent (n = 31) 2. Instructor on the left (n = 30) 3. Instructor in the middle (n = 31) 4. Instructor on the right (n = 30)	Instructor presence Instructor's position	*Not clear* Instructor on the right > instructor on the left ≈ instructor in the middle ≈ IA	*Total dwell time %* on content; IA > instructor on the left ≈ instructor on the right > instructor in the middle *Number of transitions* Instructor in the middle > instructor on the left ≈ instructor on the right
Zhang & Yang (2022)	3x2 factorial mixed design; 2 groups Video type as the within-subject IV (instructor-present vs instructor-picture vs instructor-absent) Gender as the between-subject IV (male vs female) 1. Male (n = 32) 2. Female (n = 32)	Instructor presence* Learner's gender	*Retention* Male/female + IP > instructor-picture ≈ IA	*Fixation count* On text (content); male ≈ female; IA > instructor-picture, IA > IP On picture (content); male ≈ female; instructor-picture > IA, IP ≈ IA On instructor; male ≈ female; instructor-picture > IA, IP > IA *Total dwell time* On text (content); male ≈ female; IA > instructor-picture, IA > IP On picture (content); male ≈ female; instructor-picture > IA, IP ≈ IA On instructor; male ≈ female; instructor-picture > IA, IP > IA; female + IP > male + IP
Wakefield et al. (2018)	Between-subjects; 2 groups 1. No cue (n = 24) 2. Pointing gesture (n = 26)	Hand gesture	*Not clear (combine trained and transfer problems)* Pointing gesture > no cue	*Total dwell time %* (During the strategy segment) On content; pointing gesture > no cue On instructor; no cue > pointing gesture On gesture space; pointing gesture > no cue (During the explanation segment) No significant differences between conditions were found for all AOIs
Pi et al. (2019b)	Between-subjects; 4 groups 1. No cue (n = 30) 2. Guided gaze (n = 30) 3. Pointing gesture (n = 30) 4. Guided gaze + pointing gesture (n = 30)	Attentional cue* (involving eye gaze and hand gesture)	*Retention* Guided gaze + pointing gesture ≈ pointing gesture > no cue Guided gaze + pointing gesture ≈ pointing gesture ≈ guided gaze Guided gaze ≈ no cue *Transfer* Guided gaze + pointing gesture ≈ pointing gesture > no cue Guided gaze + pointing gesture ≈ pointing gesture ≈ guided gaze Guided gaze ≈ no cue	*Time to first fixation* on the corresponding content where the instructor referred to; no cue (longer) > guided gaze ≈ pointing gesture ≈ guided gaze + pointing gesture *Total dwell time %* on the corresponding content where the instructor referred to; guided gaze + pointing gesture ≈ pointing gesture > guided gaze ≈ no cue
Wang, Pi, & Hu (2019)	2x2 mixed factorial design; 2 groups Eye gaze as the between-subjects IV (guided gaze vs no guided gaze) Knowledge type as the within-subjects IV (declarative vs procedural) 1. No cue (n = 29) 2. Guided gaze (n = 29)	Eye gaze Knowledge type	*Not clear* Guided gaze + declarative knowledge > no cue + declarative knowledge Guided gaze + procedural knowledge > no cue + procedural knowledge	*Total dwell time* On content; guided gaze > no cue On content; procedural knowledge > declarative knowledge *Number of transitions* No cue > guided gaze Declarative knowledge > procedural knowledge
Huangfu et al. (2022)	Between-subjects; 2 groups 1. Low teacher enthusiasm (n = 35) 2. High teacher enthusiasm (n = 35)	Teacher enthusiasm	*Comprehension* High enthusiasm > low enthusiasm	*Fixation count* On content; high enthusiasm ≈ low enthusiasm On instructor; high enthusiasm ≈ low enthusiasm *Total dwell time* On content; low enthusiasm > high enthusiasm On instructor; high enthusiasm > low enthusiasm

- 97% of the studies interpreted the data by linking the eye movements to overt attention.
- Instructors act as attention magnets. But when they employ techniques such as pointing gestures or guided gaze, learners focus more on the educational material.
- The relationship between eye-tracking data and learning performance is complex and varies widely, influenced by numerous factors.
- 71.4% of the reviewed studies recommend incorporating a visible instructor into video lessons, with authors advocating for the use of attentional cues (e.g., pointing gestures, guided gaze) and social cues (e.g., positive emotion, direct eye contact) during instruction.

This study was supported by the Perły Nauki grant, awarded by the Ministry of Education and Science (Nr rej.: PN/01/0232/2022)

Ministerstwo Edukacji i Nauki