Journal of Mathematical Psychology ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

An examination of parallel versus coactive processing accounts of redundant-target audiovisual signal processing

Cheng-Ta Yang[a,*], Nicholas Altieri[b], Daniel R. Little[c]

[a] Department of Psychology, National Cheng Kung University, Taiwan
[b] Department of Communication Sciences and Disorders, Idaho State University, United States
[c] Melbourne School of Psychological Sciences, The University of Melbourne, Australia

HIGHLIGHTS

- We test the redundant-target audiovisual signal processing.
- Systems Factorial Technology is adopted to test the mental architecture and capacity.
- Results are consistent with the context invariance explanation not only of the race-model inequality but also of capacity and architecture.

ARTICLE INFO

Article history:
Received 14 September 2016
Received in revised form 19 August 2017
Available online xxxx

Keywords:
Audiovisual
Race model
Redundant-target effect
System factorial technology
Context invariance

ABSTRACT

We ask whether auditory and visual signals are processed using a consistent mental architecture across variable experimental designs. It is well-known that in an auditory-visual task requiring divided attention, responses are often faster for redundant audiovisual targets compared to unisensory targets. Importantly, these redundant-target effects can theoretically be explained by several different mental architectures, which are explored in this paper. These include: independent-race models, parallel interactive models, and coactive models. Earlier results, especially redundant-target processing times which are faster than predicted by the race-model inequality (Miller, 1982), implicated coactivation as a necessary explanation of redundant-target processing. However, this explanation has been recently challenged by demonstrating that violations of the race-model inequality can be explained by violations of the context invariance assumption underlying the race-model inequality (Otto & Mamassian, 2012). We utilized *Systems Factorial Technology* (Townsend & Nozawa, 1995), regarded as a standard diagnostic tool for inferences about mental architecture, to study redundant-target audiovisual processing. Three experiments were carried out in: a discrimination task (Experiment 1), a simultaneous perceptual matching task (Experiment 2), and a delayed matching task (Experiment 3). The results provide a key set of benchmarks to which we apply several simulations that are consistent with the context invariance explanation not only of the race-model inequality but also of capacity and architecture.

1. Introduction

Multisensory processing generally refers to the ability to detect multiple signals from different sensory modalities. In many applications, multimodal stimulation increases the efficiency, measured by reaction times (RTs) of an information processing system compared to unimodal stimulation (Colonius & Diederich, 2012; Miller, 1982; Raab, 1962).

In a now classic study on multisensory processing, Miller (1982) conducted a redundant-target audiovisual detection task where participants had to simultaneously monitor auditory and visual modalities and detect the presence of an auditory target (i.e., a pure tone) or a visual target (e.g., a star presented on a computer screen). Mean RTs in the audiovisual condition (i.e., both auditory and visual targets were presented) were found to be faster than the mean RTs in the unisensory conditions (i.e., either an auditory or a visual target was presented). This well-known effect is termed the redundant-target effect (RTE) which has been a robust finding in the study of response time (Diederich & Colonius, 1987; Gondan, Niederhaus, Rösler, & Röder, 2005; Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994; Miller, 1982; Raab, 1962; Townsend & Nozawa, 1995)

Broadly speaking, the RTE has been explained by: (1) parallel models with separate channels for auditory and visual activation or information accrual (Gielen, Schmidt, & Van Den Heuvel, 1983;

* Correspondence to: No. 1, University Road, Tainan, 701, Taiwan.
E-mail address: yangct@mail.ncku.edu.tw (C. Yang).

https://doi.org/10.1016/j.jmp.2017.09.003
0022-2496/© 2017 Elsevier Inc. All rights reserved.

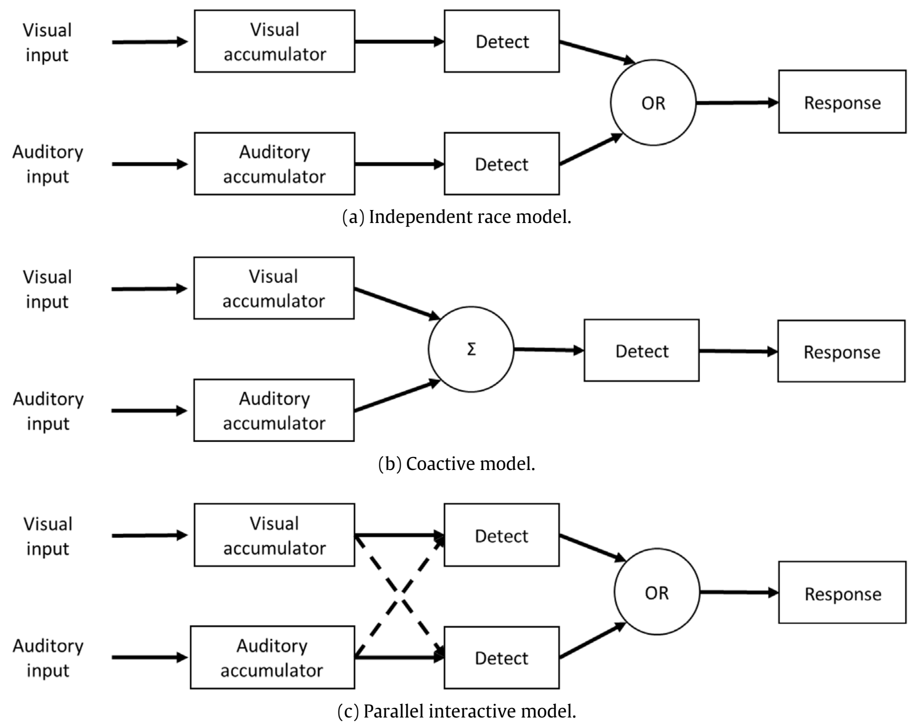

Fig. 1. Three possible mental architectures for redundant audiovisual signal processing. (A) In an independent race model, visual and auditory signals are independently processed and accumulated into separate accumulators, and outputs of the two accumulators are combined with a logical operation to trigger a response. (B) In a coactive model, signals from the two modalities are pooled together prior to decision making. (C) In a parallel interactive model, two processes are processed in parallel and non-independently.

Raab, 1962) and (2) coactive models (Diederich, 1995; Miller, 1982; Schwarz, 1989, 1994). Separate-activation models assume that auditory and visual signals are processed simultaneously and accumulated into two separate accumulators (Fig. 1(a)). A detection decision is determined by the accumulator which reaches its decision criteria first. According to the separate-activation model, the RTE occurs due to statistical facilitation from audition and vision since the minimum time of detecting the redundant target is stochastically faster than either of the single targets (Raab, 1962) .

Alternatively, the coactive model assumes that auditory and visual information originating from parallel channels are combined into a single accumulator (see Fig. 1(b)). When the summed activation reaches the decision criterion, a target is detected. According to the coactive model, RTE occurs due to the increased information accumulation rate (for instance, as dictated by a Poisson summation process; e.g., Townsend & Nozawa, 1995 and Schwarz, 1989).

1.1. Distinguishing coactive and separate-activation models

The first question that arises concerns how one might empirically distinguish parallel processing from coactive processing. One of the most general tests, historically, has been to use an inequality to determine whether the processing rate in the redundant-target condition differs from the bound predicted by a parallel processing model assuming independent channels. Miller (1982) introduced the test of the race-model inequality (RMI or Miller inequality) to distinguish between the coactive model and the separate-activation model. The RMI is expressed as:

$$P_{1,2}\left(T_{1,2} \leq t\right) \leq P_1\left(T \leq t\right) + P_2\left(T_2 \leq t\right), \quad (1)$$

where $P_{1,2}\left(T_{1,2} \leq t\right)$, $P_1\left(T_1 \leq t\right)$, and $P_2\left(T_2 \leq t\right)$represent the cumulative distribution functions (CDFs) of the audiovisual condition and the two unisensory conditions, respectively, and t represents a certain time. The inequality in (1) provides an upper bound for the performance predicted by the separate-activation model. If RMI is violated, then the separate-activation model is ruled out, and it is suggested that an alternative model, a coactive model, becomes more likely. Miller found that the RMI was violated in an audiovisual redundant-target detection task and in an audiovisual discrimination task.

There are several reasons to interpret this result in favor of coactive processing with a degree of caution. First, a critical assumption of the RMI is that the processing times for the single channels have the same distribution as the marginal times for each channel when both targets are presented jointly (i.e., the redundant-target condition; (Colonius, 1990)). That is:

$$P_{1,2}\left(T_1 \leq t\right) = P_1\left(T_1 \leq t\right) \text{ and } P_{1,2}\left(T_2 \leq t\right) = P_2\left(T_2 \leq t\right). \quad (2)$$

This assumption is known as *context invariance* (CI, or context independence, though we prefer the former term to avoid confusion with the term statistical independence; (Ashby & Townsend, 1986; Colonius, 1986, 1990; Colonius & Townsend, 1997)). This assumption implies that the distribution of the processing time for a channel (i.e., audition) under unimodal stimulation should be identical to the distribution of the processing time for that channel under audiovisual stimulation. Unfortunately, this assumption does not necessarily hold in audiovisual detection.

Two ways in which the assumption of context invariance might fail are if the mean or the variance of one of the channels increases or decreases when presented together with the other channel compared to when presented alone. If the mean RT of a channel decreased or if the variances of a channel increased when presented as part of a redundant target, then this would allow a separate-activation model to violate the RMI (Otto & Mamassian, 2016). Otto and Mamassian (2012) found that data from an audiovisual redundant-target detection task was best fit by a model that allowed for increased noise in both channels when both