RESEARCH REPORT

Gaze and hand behaviors during haptic abilities testing—An update to multimedia learning theory

Michelle A. Sveistrup[1] | Jean Langlois[2] | Timothy D. Wilson[1]

[1]The Corps for Research of Instructional and Perceptual Technologies (CRIPT) Laboratory, Department of Anatomy and Cell Biology, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

[2]Department of Emergency Medicine, CIUSSS de l'Estrie – Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, Quebec, Canada

Correspondence
Dr. Timothy D. Wilson, The CRIPT Laboratory, Department of Anatomy and Cell Biology, Medical Sciences Building, Rm 443, Schulich School of Medicine and Dentistry, Western University, London, ON N6A 5C1, Canada.
Email: tim.wilson@uwo.ca

Abstract

The Cognitive Theory of Multimedia Learning (CTML) suggests humans learn through visual and auditory sensory channels. Haptics represent a third channel within CTML and a missing component for experiential learning. The objective was to measure visual and haptic behaviors during spatial tasks. The haptic abilities test (HAT) quantifies results in several realms, accuracy, time, and strategy. The HAT was completed under three sensory conditions using sight (S), haptics (H), and sight with haptics (SH). Subjects ($n=22$, 13 females (F), 20–28 years) completed the MRT (10.6 ± 5.0, mean $\pm SD$) and were classified as high or low spatial abilities scores with respect to mean MRT: high spatial abilities (HSA) ($n=12$, 6F, MRT$=13.7 \pm 3.0$), and low spatial ability (LSA) groups ($n=10$, 7F, MRT$=5.6 \pm 2.0$). Video recordings gaze and hand behaviors were compared between HSA and LSA groups across HAT conditions. The HSA group spent less time fixating on mirrored objects, an erroneous answer option, of HAT compared to the LSA group (11.0 ± 4.7 vs. 17.8 ± 7.3 s, $p=0.020$) in S conditions. In haptic conditions, HSA utilized a hand–object interaction strategy characterized as palpation, significantly less than the LSA group (23.2 ± 16.0 vs. 43.1 ± 21.5 percent, $p=0.022$). Before this study, it was unclear whether haptic sensory inputs appended to the mental schema models of the CTML. These data suggest that if spatial abilities are challenged, LSA persons both benefit and utilize strategies beyond the classic CTML framework by using their hands as a third input channel. This data suggest haptic behaviors offer a third type of sensory memory resulting in improved cognitive performance.

KEYWORDS
cognitive load, haptic abilities, learning, spatial abilities, strategies

INTRODUCTION

In modern learning environments, the cognitive theory of multimedia learning (CTML) provides a model where cognitive loads, the mental workload required to learn, are the mediation of the appropriate and constituent loads involved. In particular, extraneous loads generated by the learning environment are to be minimized.[1] At least three types of cognitive load may be considered with any learning task: intrinsic, extrinsic, and germane loads.[1,2] The CTML is based on Paivio's Dual-Coding Theory (DCT)[3,4] and three

assumptions therein.[3,4] First, humans possess separate cognitive processing systems through two sensory channels: one visual and one auditory.[5] Second, each channel has a finite processing capacity.[5] Finally, learning involves cognitive processing to build mental schema between pictorial and verbal representations.[5] Initially organized in the learner's working memory, schemas are constantly updated and appended to long-term memory enabling recognition, categorization, and actions toward problems presented.[6,7] Mental respresentations are the resulting effects of encoding information into one's long-term memory.[8]

Vision is the principal sense used by sighted individuals to gather information about their surroundings while haptic perception offers different elemental information of one's immediate environment,[9] especially if vision is obscured or unclear. The constructs of what haptics entail differ with fields of application and study. In general, haptics are considered the ability of reaching with the arm and hand to touch, grasp, and potentially manipulate an object. The hand tasks are usually defined in terms of the action performed by the hand with, or to, the object (e.g., feeling a texture, grasping a scalpel, or writing with a pen).[10] Therefore, haptic perception is defined as active touch-based sensory interaction with physical objects.[11]

To accurately manipulate one's mental representations of shape, orientation, and spatial relationships, and to derive accurate meaning of this knowledge, individuals rely on spatial abilities (SA).[12] Human SA are described as a suite of cognitive capacities used to apprehend, remember, generate, and manipulate mental representations of objects, diagrams, maps and spatial relations to problem-solve within these contexts.[12] In sighted individuals, SA are predominantly driven by vision and are thought to be an important factor contributing to the perception and interaction with visual surroundings.[13] Reports suggest SA are a predictor of success in science, technology, engineering, and mathematics (STEM) disciplines.[14] In anatomy, high SA (HSA) appear to aid visualization of various sections and planes, translation and rotation of anatomical objects, leading to improved practical anatomy knowledge assessment.[15–18]

Through common tests, SA are quantitatively represented.[19,20] Although the tests are numerous, SA are often based on Shepard and Metzler's approach.[12] Termed the Mental Rotations Test (MRT), it requires visualizing a drawn geometric object and transforming a mental representation of the object to ultimately match it to its identical but rotated like-pair rather than mirrored images.[12] The most common and modern used test was created by redrawing the original questions in the MRT by Vandenberg and Kuse[20,21] and testing in electronic interfaces under time duress[15] (Figure 1A).

Literature suggests SA is observable in eye fixations.[23] Fixations correspond to maintaining ocular gaze for a minimum time between the dynamic components of visual attention, saccades.[24] Saccades are rapid eye movements between fixations and serve to bring an
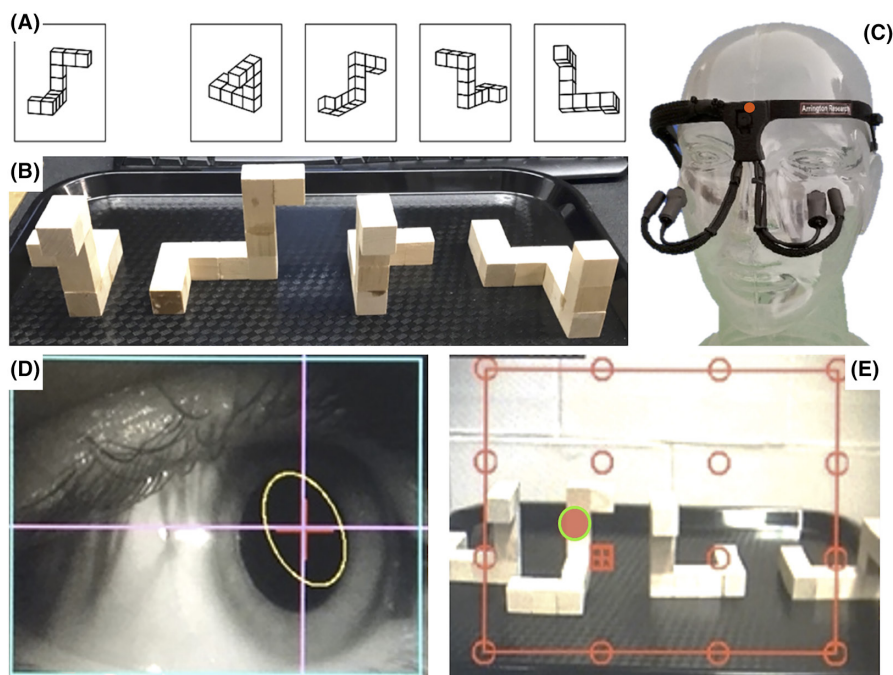

FIGURE 1 Panel A—The redrawn MRT used in electronic interfaces. Participants must determine which two of the four objects on the right are a rotated version of the exemplar on the left. Panel B—Example question for the HAT. The exemplar is located on the left followed by the correct match, mirrored incorrect match and incorrect match. Panel C—The head-free eye-tracker with an outfacing scene camera (indicated by red circle) at bridge of nose to collect field of view information, an inward (eye) facing camera and infrared (IR) light on adjustable arms, to record eye movements. Panel D—Gaze-tracking software recording of subject's pupil under IR light and tracking target. Panel E—Eye tracking scene camera field of view recording of the HAT with pupillary position (indicated by red circle with green outline). Gaze tracking was measured in S (sighted) and SH (sighted with haptics) conditions of the HAT. Panels A and B are used with permission.[22]

object or region of interest into focus on the fovea.[25] During saccades, visual inputs are unattended as they are smeared across the retina and severely degraded thus not offering useful visual information.[26] Once a saccade is completed, however, the projection of the visual field onto the retina has undergone a large-scale spatial transformation.[26] Fixations usually have a duration of ~250 ms[24] and it is primarily during fixation that visual processing occurs.[27]

Differences in fixations were identified in individuals with high and low SA (H/LSA) in a study exploring the relationship between eye movement and spatial reasoning when answering questions on different tasks sampling mental rotation abilities.[28] The HSA subjects respond faster and more accurately than their LSA counterparts and time to responses increases as testing shapes diverge in angular disparity. The authors suggest that LSA individuals undertake more cognitive work on the same challenge suggesting lower efficiency in these tasks. Interpreted through an information processing framework, this is increased cognitive load. Capitalizing on potentially innate attentional behavioral differences between H/LSA, Roach et al. explored discreet mechanistic approaches to better delineate how SA is expressed in selecting salient aspects within images.[29] Using an electronic MRT in concert with eye tracking they concluded that HSA and LSA groups view identical images with different observational strategies, leading often to different conclusions.[29]

Current literature provides far less information about haptic abilities (HA) compared to spatial abilities (SA), yet the field is rapidly advancing.[30] Human HA are considered a vital aspect of the human perceptual system relying on touch sensitivity and the integration of partial tactile information to form mental representations of objects.[30] Numerous studies highlight the significance of HA in object recognition,[30–33] yet little mechanistic investigation[34] has informed learning theories or how haptics are especially important in allied health education. The incorporation of "hands-on" learning is one of the earliest forms of sensory inputs any experiential learning, and anatomy laboratories, trade schools, and physical examinations rely heavily on developing adept skills therein.

As noted primarily with high SA, the development of one's HA may also offer advantages in STEM disciplines, visual disability instruction, and language learning.[35–39] The use of haptic technology in education has proven to be beneficial, leading to improved learning outcomes in various fields.[40] In the realm of medical training, appropriate haptic input can provide students with a lifelike experience in surgical training, resulting in enhanced performance during laparoscopic surgery simulations.[35–38,41] Furthermore, haptic input has proven useful for veterinary students in gaining a better conceptualization of bovine abdominal anatomy and its three-dimensional (3D) visualization.[42] The mechanisms by which immersive haptic experiences result in improved learning outcomes occurs presumably through the moderation of cognitive load, increasing students' germane load, and/or decreasing the extrinsic loads during learning activities have received little research attention. The enhancements achieved suggest an increased interest and curiosity in the learning materials that engage learners by improving the ability to connect with the topic and constructing mental representations of

abstract concepts through hands-on experiences.[40,43] When handling anatomical structures[18,44] and learning technical skills in health care,[45–47] vision can only provide the learner with partial information[41] as pressures, torque, surface tensions, textures, and confinement cannot solely be sensed visually. Consequently, the process of learning such tasks involves, at least partially, haptic exploration[41] and the integration of sensory inputs.

It is currently unknown whether haptic sensory inputs append to the mental schema representing images constructed via sensory channels of the CTML or exist as unique haptic-generated mental representations within the working memory. Using a functional magnetic resonance imaging (fMRI) haptic repetition paradigm, Snow et al. found that visually defined areas in the occipital and temporal cortex were involved in analyzing the shape of touched objects.[48] Their findings suggest that shape processing via touch engages many of the same neural mechanisms as visual object recognition. The engagement of the primary visual cortex demonstrates that "visual sensory" areas are engaged during the haptic exploration of object shape. In the absence of concurrent shape-related visual input, shape cues processed within visual circuits may be relayed back to somatosensory and motor areas to guide ongoing exploration.[48] Chow et al. used haptic recognition of shape surfaces and believed that the exploratory movements subjects adopted, thus, the type of information each subject gathered for object discrimination could be a critical difference between the two haptic tests they created.[49] The authors did not test visual object recognition, nor did they consider whether SA interacts therein despite sharing the aforementioned sensory overlap; nonetheless, they conclude that haptic object recognition ability is not purely related to visual recognition, but more likely related to the processing of the underlying task.[49] Thus, haptics may represent a third type of sensory memory within the context of CTML model.[1] If haptic sensory "memory," the terminology of the model, integrates into working memory akin to the auditory and visual channel memory, the use of haptic inputs could improve learning by distributing cognitive loads across differing sensory channels even further. Here, one's sense of touch contributes to formulations of haptic schema that integrate in working memory. The integration enables the learner to select from a palate of sensory inputs that contribute to higher accuracy in the learning environment and better long-term memory (LTM). This begs the question of whether an ability that is dominated visually like SA is shared with, or compensated for, by HA. If an individual has a limited capacity to apprehend salient aspects within the learning materials with one sensory modality, as seen in low SA individuals (LSA), then rapid and seamless sensory switching to alternative modalities is imperative for consistent learner comprehension.

The purpose of the current study was to measure gaze and hand-to-object interactions in HSA and LSA individuals during a haptic abilities test (HAT). It was hypothesized that HSA and LSA individuals would demonstrate similar gaze and hand behaviors when allowed to use both sensory modalities, but the behaviors would differ when either vision or haptics was removed, and individuals relied solely on singular sensory inputs to complete the HAT.

MATERIALS AND METHODS

Participants

A total of 22 students (nine males, 13 females; age: 19–28 years) enrolled in undergraduate degrees at the University of Western Ontario, London, Ontario, Canada with normal, or corrected vision, participated in this prospective cohort study. The methodologies were approved by the institution's Research Ethics Board (ID: 118803). Participants were recruited via class or online announcements within the institution's learning management system from various departments across the campus. All participants provided written informed consent. Performance data pertaining to the MRT and HAT scores from these participants are reported in a prior manuscript by our lab.[22] The results in this manuscript describe new data from gaze-tracking and hand–object behaviors, foundational to the perception to action underpinning previously reported HAT scores.

EXPERIMENTAL PROTOCOL

The experiment was divided into two parts. Part I is identical to the methods defined previously[22] while Part II explores behaviors underlying HAT scores. Each participant was compensated with a CAD$20 sum in the form of a campus food card.

Part I: Spatial abilities testing

Participants completed a screening questionnaire to identify their primary field of study, sex, age, handedness, previous SA testing, and video game frequency. Consenting participants then completed an online MRT originally designed by Vandenberg and Kuse[20] and later redrawn[21] (Figure 1A). The test consists of 24 multiple-choice-style questions. Each subset of 12 questions had a time limit of 3 min and was separated by a 3-min break. Participants were asked to identify the two correct, but rotated, images of the exemplar structure to obtain one point. The maximum score was 24 points. Individuals with higher SA complete this test in less time and with greater accuracy than those with lower SA.[16] Therefore, test scores obtained on the online MRT enabled the categorization of SA into HSA or LSA participant groups. For this study, participants were categorized as HSA or LSA depending on whether their score on the MRT was above or below the average score of all participants.

Part II: Haptic abilities testing

Participants underwent a pre-assessment of dexterity, the Purdue Pegboard Test (Lafayette Instrument Inc., Lafayette, IN, USA). This assessment is used to measure unimanual and bimanual finger and hand dexterity.[50] Under a 30 s time constraint, participants placed as many small metal pins, washers, and collars systematically on a peg board. The test is undertaken with each hand and bimanually. The total number of pin/washer/collars form an assembly score. Scores were noted using the Purdue Pegboard scoring application's standardized administration (Lafayette Instrument Inc., Lafayette, IN, USA).

Next, participants performed an object manipulation test, termed the Haptic Abilities Test (HAT).[22] The HAT quantifies an individual's HA using 3D objects. This test requires the decision of orientations of handheld 3D wooden objects (~10 cm³). The 3D wooden objects mimicked the original MRT objects drawn by Shepard and Metzler[12] and were previously reported in Sveistrup et al.[22] (Figure 1B). Each unique shape identified in the online two-dimensional (2D) MRT of part I was replicated to form the 3D objects used in the HAT. While the MRT positions objects in 3D space at a variety of orthogonal orientations, the HAT objects could only be positioned on a flat surface. Thus, the incident viewing angle differed slightly for participants based partially on their height and partially on their chosen posture to view the HAT objects. Each question of the HAT contains four wooden objects: the exemplar and three possible answers. The correct answer for each question consists of a rotated matching object, and two erroneous choices consisting of a mirrored image of the exemplar and one incorrectly shaped object. The HAT differs from the MRT where four possible answer options are possible. The degree of angular differences between the exemplar and the correct match was altered for every question; thus, the correct match was consistently rotated at 90°, 180°, and 270° from the exemplar object.

The HAT is undertaken three times under a different, and randomized, sensory conditions. A three to five minute break was given between conditions. The conditions include a sighted only (S) condition, where participants relied solely on their sense of vision; a sighted + haptics (SH) condition; and a haptic only (H) condition where a curtain obscured views of the HAT objects requiring haptic examination alone. In all conditions, answers were indicated by holding, pointing, touching, and/or verbalizing their answer. The order of the sensory conditions were randomized using a random number generator between participants to reduce the effects of testing. While the MRT imposed a 3-min time to complete each testing battery, the HAT test did not impose time limits a priori. The HAT instructions remind participants that the best scores are achieved by the fastest and most accurate answers. No feedback or indications of accuracy are given to participants.

Each of the HAT sensory conditions consisted of 15 questions. Within each condition, five questions were repeated to sample repeatability of participants responses and calculate reliability indices. Participants did not report awareness of any repetition. The question order and the block orientations therein were identical for all three conditions enabling comparison between individuals and conditions. For scoring, one point was allocated to participants who chose the rotated matching object, and zero points were allocated for choosing the mirrored or the incorrect match resulting in a maximum of 15 points per condition.

INSTRUMENTATION

Gaze tracking

Gaze parameters were measured with a head-free eye-tracking headset (Arrington Eye-Tracker – Arrington Research Inc., Scottsdale, AZ, USA), resembling a pair of lightweight lens-less glasses containing two cameras and one infrared light source facing the eye, for the entire duration of the HAT protocol (Figure 1C). The first camera is an outward-facing scene camera integrated into the frames of the headset and positioned at the bridge of the nose. It records the participant's field of view at a sampling frequency of 60 Hz. An inward-facing infrared eye camera records eye behaviors and is illuminated by an infrared light source (Figure 1D). The gaze-tracking software enables both simultaneous and offline detection of pupillary fixations, locations, and pupillary diameter at a sampling frequency of 60 Hz. This technique affords insights into cognitive processing at the time of fixation and accompanying attention underlying behavior through the eye–mind relationship.[29]

The eye tracking software (ViewPoint, Arrington Research Inc., Scottsdale, AZ, USA) utilizes a synchronous playback and recording system to overlay pupillary coordinates (*x, y*) fixation locations on the video scenes collected from the scene camera (Figure 1E). Measuring pupillary behaviors during fixations affords the recording of the number of fixations, fixation durations, and pupillary diameter during fixation during each question, in each condition, of the HAT protocol where the vision is allowed, sighted (S) and sighted + haptics (SH) conditions. The spatial and temporal information enables the characterization of a participant's gaze behaviors and serves as an index of visual attention. Since visual processing only occurs during the fixations,[27] both the location within the visual field and temporal parameters regarding that particular fixation were derived with question-by-question granularity and calculated to compare across individuals on a question, condition, and ability basis.

Temporal parameters of gaze

The minimum fixation duration was determined a priori at 250 ms. The temporal parameters of gaze are represented on a question-by-question basis in the S and SH conditions. The fixation number (FN) was defined as the total number of times the pupil fixated on a specific location while average fixation durations (AFD) were calculated as the mean duration of all fixations occurring in each question (s). The fixation rate (FR) indicates how many fixations were undertaken before the subject answered; therefore, the FN is divided by the reaction time (RT) per question (fixations/s).

Spatial parameters of gaze

Gaze locations on the HAT objects were determined through analysis of scene camera recordings of the participant's field of view with an overlaid recording of the pupillary *x–y* location within that visual space (Figure 1E). Through characterization of gaze location, assessment of visual attention is derived by quantifying the amount of fixation time spent observing each of the four objects within each question of the HAT during the S and SH conditions. In the H condition, no pertinent gaze information is present.

In all conditions of the HAT, pupillary diameter (PØ) was recorded by the inward-facing camera. The PØ was measured as an indirect measure of stress. Participants' changes in pupillary diameter (ΔPØ) were expressed relative to the average PØ measured during a minimum 20 s baseline and were calculated on a question-by-question basis to infer participants' cognitive loads as the pupils dilate[51] with differing sensory inputs in each of the HAT conditions.

Measurements of hand–object manipulation behaviors and strategies

In conditions affording haptic somatosensory input of test objects (SH and H), the hand–object manipulations were recorded using an external video recorder mounted in a stationary, superior, and lateral position (Figure 2). The nature and time of interactions between participants' hands and objects were determined offline. Reviewing video on a question-by-question basis enabled identification hand–object interaction. The approach draws on grounded theory methodology using thematic codes in qualitative research. Here, instead of coding participant verbal feedback, hand–object behaviors are characterized and categorized. These video-based observations are recorded temporally (time-stamped), the hand–object behaviors are interpreted by an observer and recorded as codes of stereotypic gestures. As codes accumulate, similar hand–object interactions may be grouped as they share commonalities. Therein, a strategy classification rubric emerges, similar to coding

Curtain Obstructing Participant's View of Objects

FIGURE 2 Static image from video recordings used to determine hand–object interactions. In this image, the subject is reaching through the curtain to interact with HAT objects in the H condition.

in qualitative approaches using participant conversations.[52,53] As the number of meaningful, task-related, hand–object interactions are finite and likely stereotypical,[54,55] a coding event called saturation will occur. Contextually, saturation indicates that no new hand–object strategies are expressed.[52,53] The rubric of strategy codes thus describes all meaningful data. The method affords several layers of analysis in determining what and how each of the strategies are used: the determination of hand–object interaction (strategy), the frequency of use of the strategies identified, and the accumulated time of each strategy within a question, all become quantifiable events. Dividing the accumulated amount of time individuals use each strategy by question response time, the percent of use of each strategy is calculated on a question-by-question basis. The use of the classification rubric enabled comparisons of behaviors utilized during the two HAT conditions where hands may be used to achieve the tasks (SH and H) and any strategy-based differences expressed by HSA and LSA individuals on a question-by-question basis.

Data analysis

Statistical analysis was undertaken with SPSS Statistics for Windows (Release 27.0, IBM Corp, Armonk, NY, USA) while visual plots were created using Microsoft Office Excel (Version 16.54, 2021). A Shapiro–Wilk test was used for each dependent variable across HAT conditions to determine whether data were normally distributed. Where appropriate, descriptive parametric data were reported as means $\pm SD$ while inferential data were deemed significant with an alpha value set a priori at 0.05. In some instances, the data did not follow normal distributions, these non-parametric data were reported as median values with lower and upper quartiles (Q1, Q3).

Intra-rater correlation coefficients, a measure of reliability, were calculated using the total time accumulated for each strategy for each of the questions 1–5 and 11–15. These questions were identical across all subjects. For inter-rater reliability, correlation coefficient was undertaken using two independent raters who quantified the hand–object interaction strategies (frequency and time) expressed with a random subset of four individuals (2F/2M) from the overall pool.

Participant characteristics

Discrete variables such as the number of participants, sex, study discipline, previous SA testing, handedness, and video game frequency were reported as frequencies (n). Participants were categorized into HSA and LSA groups based on their MRT scores. Independent samples t-tests were used to compare the independent variables of age, video game frequency, MRT, and manual dexterity scores between HSA and LSA individuals. Fisher's exact test was used to compare the discrete variables between HSA and LSA groups.

Gaze behavior

A Wilcoxon signed-rank test was used to compare fixation number (FN) and average fixation duration (AFD) between the S and SH conditions, whereas a paired samples t-test was used to compare fixation rate (FR) between the S and SH conditions in all participants. A Mann–Whitney U-test was used to identify differences in FN and AFD between HSA and LSA individuals in the S and SH conditions and an independent samples t-test was used to compare FR between HSA and LSA participants in the S and SH conditions of the HAT.

Gaze location behavior, used to assess visual attention, was defined as the percentage of time fixating on each of the four HAT. The location of each fixation was noted as one of the four HAT objects (exemplar, correct match, incorrect match, or mirrored match) per question per condition. The cumulative time within each question fixating on a certain HAT object was totaled to obtain a total percentage of time spent fixating on each HAT object. The difference in gaze location behavior per object was compared using a one-way repeated measures analysis of variances ANOVA in the S condition and a Kruskal–Wallis testing in the SH condition. The gaze location data were compared between the S and SH conditions using a Friedman test. Gaze location differences were compared between HSA and LSA individuals in the S and SH conditions using an independent samples t-test and Wilcoxon signed-rank test respectively.

Pupillometry

A one-way repeated measures ANOVA was used to compare the percentage change in pupillary diameter (ΔPØ) between the S, SH and H conditions of the HAT in all participants. Comparison of ΔPØ between HSA and LSA participants in the S, SH, and H conditions of the HAT were analyzed using a split-plot ANOVA (factor 1 = SA, 2 levels, i.e., HSA vs. LSA, factor 2 = condition, 3 levels, i.e., S vs. SH vs. H).

Haptic time

Video from the external video recorder was used to determine the objects garnering the greatest haptic attention in the SH and H condition in HAT. The general approach of haptic time (HT) represents the relative time spent (%) touching each of the HAT objects. The HT identifies the most haptically salient object using dominant and non-dominant hands. The comparisons of HT per object of the HAT were compared across all questions using a Kruskal–Wallis test, whereas any difference in the HT between HSA and LSA was compared using Mann–Whitney U-test.

Haptic strategies

The haptic strategies during the SH and H conditions were identified post hoc. Comparison of each strategy usage between HSA and

LSA participants was analyzed using an independent *t*-test for the normally distributed strategies and a Mann–Whitney *U*-test for the non-parametric strategy data. A comparison of strategy utilization between groups (H/LSA) and across conditions (SH/H) was analyzed using a Friedman test.

RESULTS

Participant characteristics

Table 1 presents the characteristics of the participants from this study. Of the total participant pool, 13 (59.1%) were females. There were no significant differences in MRT scores based on sex ($p = 0.283$). Participants identified themselves as either right-handed ($n = 21$) or left-handed ($n = 1$). The average MRT score was 10.6 ± 5.0. Participants were categorized into HSA ($n = 12$) or LSA ($n = 10$) groups based on whether their MRT score was above or below the average score (13.7 ± 3.0 vs. 5.6 ± 2.0, $p < 0.0001$). Manual dexterity scores were within the normal range and did not differ significantly between the HSA and LSA groups (136 ± 11 vs. 130 ± 12, $p = 0.281$). There was no significant difference in age between the HSA and LSA groups (21 ± 2 vs. 21 ± 3 years, $p = 0.872$), nor in video game play frequency (6 ± 6 vs. 4 ± 8 h/week, $p = 0.107$). Three participants (14%) were enrolled in social sciences disciplines.

Attention—gaze behaviors

Gaze-related parameters were obtained for 20 participants as two were unusable due to poor video quality (HSA: $n = 12$; LSA: $n = 8$). The fixation rates (FR) on a question-by-question basis were normally

TABLE 1 Participants characteristics.

	HSA	LSA	*p*-Value
n	12	10	n/a
MRT score	14±3	6±2	<0.0001
Age (years)	21±2	21±3	0.872
Female	6.0	7.0	0.415
Male	6.0	3.0	0.415
Right-handed	10.0	9.0	0.481
Left-handed	0.0	1.0	0.481
Ambidextrous	2.0	0.0	0.481
STEM disciplines	11.0	8.0	0.571
Non-STEM disciplines	1.0	2.0	0.571
Video game frequency (h/week)	6±6	4±8	0.107
Previous SA testing	0.0	0.0	1.000
Manual dexterity	136±11	130±12	0.281

Note: Data reported as means ± standard deviation.
Abbreviations: HSA, high spatial abilities; LSA, low spatial abilities.

distributed; however, the number of fixations (FN) and average fixation durations (AFD) were not. The gaze location data were normally distributed for the S condition and were not for the SH condition.

In the S and SH conditions, there were no significant differences in mean FN (6.1 ± 3.0 vs. 6.3 ± 3.4, $p = 0.681$) (Figure 3A), AFD (1.3 ± 0.6 vs. 1.3 ± 0.7 s, $p = 0.218$) (Figure 3B), and FR (0.9 ± 0.2 vs. 1.0 ± 0.2 fixations/s, $p = 0.103$) (Figure 3C). No significant differences in mean FN were detected between HSA and LSA individuals in either the S and SH conditions (S: 5.9 ± 2.4 vs. 6.3 ± 3.9, $p = 0.784$; SH: 5.3 ± 1.6 vs. 7.8 ± 4.8, $p = 0.238$), AFD (S: 1.4 ± 0.7 vs. 1.1 ± 0.5 s, $p = 0.734$; SH: 1.45 ± 0.9 vs. 1.0 ± 0.4 s, $p = 0.343$) and FR (S: 0.9 ± 0.2 vs. 0.9 ± 0.2 fixations/s, $p = 0.916$; SH: 1.0 ± 0.3 vs. 1.0 ± 0.2 fixations/s, $p = 0.797$).
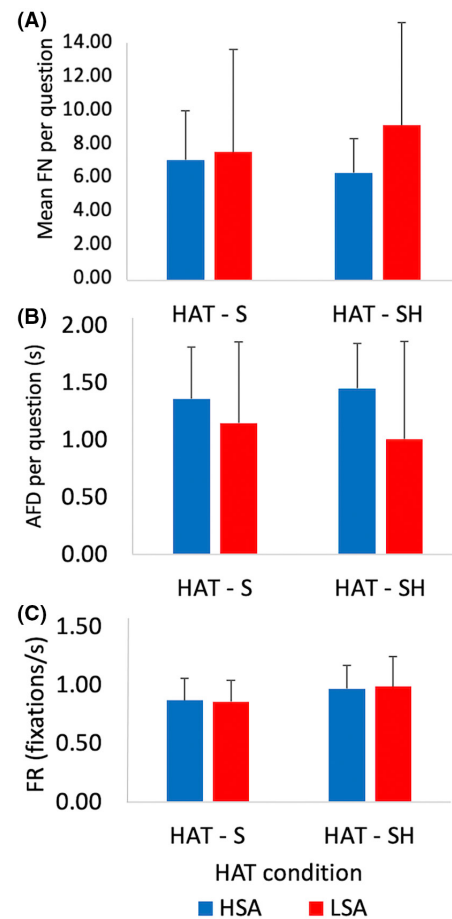

FIGURE 3 Participant's visual attention during the HAT represented by gaze fixations during sighted S and SH conditions in HSA and LSA individuals ($n = 20$). Data are represented as per question means ±1*SD*. Top Panel A: Average fixation number (FN) indicating how often participant's eyes remain stationary per question. Middle Panel B: Average fixation duration (AFD) indicates how long the eyes remained stationary during each question. Bottom Panel C: Average fixation rate (FR) indicates how often individuals moved their eyes from location to location within their field of view. No significant differences between groups or conditions were found. HAT, haptic abilities test; HSA, high spatial abilities, LSA, low spatial abilities; *n*, frequency; s, seconds; S, sighted; *SD*, standard deviation; SH, sighted haptic.

Gaze location was analyzed as the percent of participants' individual response time/question spent fixating on each of the HAT objects in the S and SH conditions. In the S condition, the correctly matched object was fixated upon significantly more than the exemplar, incorrect match, or mirrored match ($p<0.001$) (Figure 4A). In the SH condition, the exemplar and the correct match were fixated on significantly more than the incorrect or the mirrored match ($p<0.001$) (Figure 4B). Further, in the S condition, the time spent fixating on the mirrored match was significantly higher in the LSA group compared to the HSA group (17.8 ± 7.3 vs. 11.0 ± 4.7, $p=0.020$) (Figure 4A). However, there were no significant differences in the time spent fixating the exemplar (31.0 ± 10.7 vs. 30.4 ± 10.0, $p=0.910$), the correct match (45.4 ± 10.4 vs. 38.6 ± 9.6, $p=0.159$) and the incorrect match (12.7 ± 5.5 vs. 13.2 ± 7.0, $p=0.868$) between HSA and LSA groups in the S condition (Figure 4A). In the SH condition, there were no significant differences in the time spent fixating on the exemplar for H/LSA (37.4 ± 14.3 vs. 40.4 ± 18.3, $p=0.681$), the correct match (35.9 ± 5.7 vs. 32.2 ± 12.7, $p=0.384$), the incorrect match (12.4 ± 9.8 vs. 12.6 ± 9.1, $p=0.849$), and mirrored match (14.7 ± 7.3 vs. 14.8 ± 6.3, $p=0.987$) between HSA and LSA groups (Figure 4B).

Pupillometry

Pupillometry data were collected in the three conditions of the HAT in all 22 participants. The pupillary diameter (PØ) was reported as the linear distances measured at the maximum horizontal and vertical axis at the edges of the pupil. The percent ΔPØ significantly decreased in the H condition in both HSA and LSA groups (effect of group: $p=0.893$, effect of condition: $p=0.049$, Figure 5).

Attention—haptic behaviors

Haptic time

The percentage (%) of time spent touching and/or manipulating HAT objects and the haptic strategies used during the SH and H conditions of the HAT were compared in all right-handed individuals ($n=21$). Between HSA and LSA participants, in the SH and H conditions, there were no significant differences in the percentage of time spent manipulating the objects in the HAT. However, granular
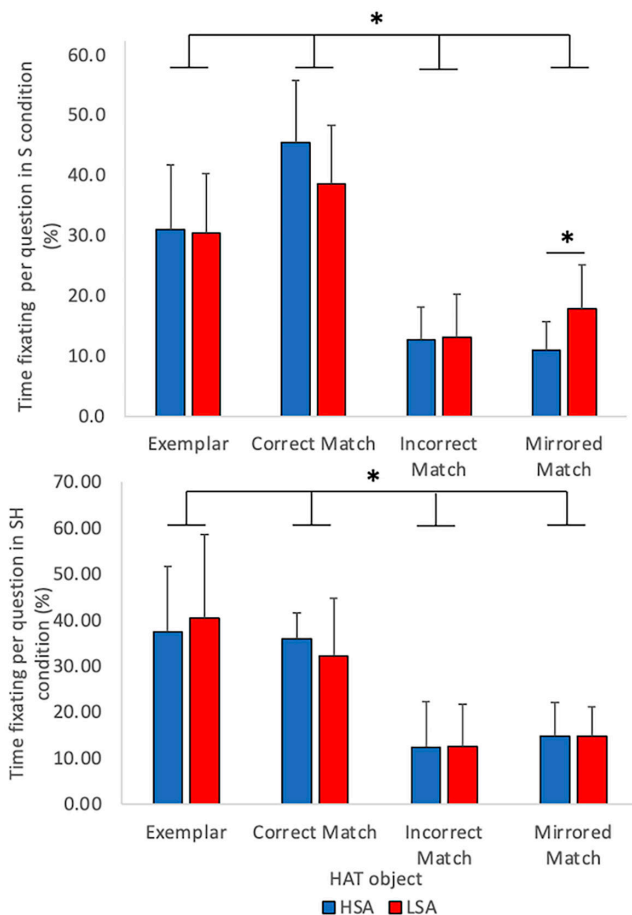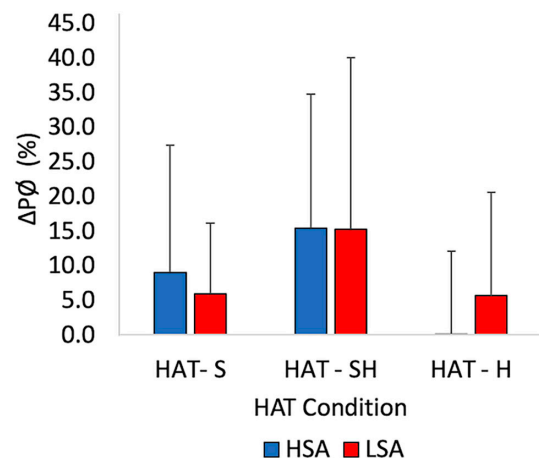

FIGURE 4 Gaze fixation location represented as percent of time to response on each of the 3D objects in the HAT conditions using vision. Each panel illustrates times of HSA and LSA participants ($n=22$). Top Panel—The average percent of time fixating on each of the four HAT objects in the S condition. The correct match object was attended to more than the incorrect or mirrored match ($p<0.001$). The mirrored match held the attention longer for LSA versus HSA individual ($p=0.020$). Bottom Panel—Time fixating on each of the four HAT objects in the SH condition in HSA and LSA individuals ($n=22$). The exemplar and correct matched were significantly more attended to than the incorrect and mirrored match (*$p<0.001$). Error bars indicated $\pm1SD$. 3D, three-dimensional; HAT, haptic abilities test; HSA, high spatial abilities, LSA, low spatial abilities; n, frequency; p, p-value; S, sighted; SD, standard deviation; SH, sighted haptic; %, percent.

FIGURE 5 Mean percent change in pupillary diameter (ΔPØ) from participants' baseline diameter in the S, SH, and H sensory conditions of the HAT for HSA and LSA groups ($n=22$). No differences between SA groups are present but the main effects of sensory conditions ($p=0.049$) are present. Error bars indicate $\pm1SD$. H, haptic; HAT, haptic abilities test; HSA, high spatial abilities, LSA, low spatial abilities; n, frequency; p, p-value; S, sighted; SD, standard deviation; SH, sighted haptic; %, percent.

assessments of the percent of response time manipulating the correct, incorrect, and mirrored objects with the dominant hand were lower in the H condition compared to the SH condition for both spatial ability groups. In the HSA group, object manipulation time as a percent of each question's RT in H versus SH (Correct: 28.6%±12.4% vs. 40.2%±6.6%, $p=0.016$; Incorrect: 7.4%±6.5% vs. 21.4%±6.0%, $p<0.001$; Mirrored: 5.8%±5.2% vs. 21.1%±5.5%, $p<0.001$) and for LSA individuals (Correct: 23.3%±14.0% vs. 36.7%±5.3%, $p=0.006$; Incorrect: 10.4%±7.3% vs. 22.22%±4.6%, $p<0.001$; Mirrored: 9.4%±10.8% vs. 23.6%±7.9%, $p=0.006$) (Figure 6). The exemplar object was manipulated significantly less in the SH condition compared to the H condition in the HSA group (0.7%±1.9% vs. 5.7%±6.6%, $p=0.026$) (Figure 6). The exemplar in the non-dominant hand was manipulated significantly less in the SH condition compared to the H condition in both groups (HSA: 19.0%±31.4% vs. 91.9%±14.7%, $p<0.001$; LSA: 26.8%±40.6% vs. 92.7%±17.1%, $p=0.004$) (Figure 6). Figure 2 demonstrates a still picture taken from video footage.

Haptic strategies

The consistency of the video-based detection and documentation of individual, question-by-question strategy was reliable and repeatable. Comparisons of repeated questions of the HAT (Q1-5 and Q11-15), enabled comparison of intra-rater consistency. The intra-rater Pearson correlation coefficient was ($r=0.84$) indicating high levels of internal consistency. To assess inter-rater measures of reliability of

the approach, a subset of 2F and 2M participants were re-analyzed with two independent raters using the same method but with all 15 HAT questions in the haptic condition. The inter-rater Pearson correlation coefficient was calculated ($r=0.79$) indicating a high level of agreement between the raters.

To delineate haptic exploration beyond the description of object touch frequency and time, a hand–object strategy rubric containing common strategies was created *post-hoc*. From the strategy saturation method employed, eight haptic strategies emerged from this data. The classification of contextually based strategies include dynamic tracing, static hold, palpation, object lift, rotation, object matching, revisiting objects, and using an exclusion method (Table 2). The dynamic tracing, static hold, palpation, and lifting strategies were analyzed individually as a proportion of the cumulative time they occurred per question in the SH and H conditions. The use of rotation, matching, revisiting, and exclusion strategies was categorized in a binary manner, based on frequency of use, per question.

Overall, HSA individuals used the palpation strategy proportionally less than LSA participants (23.2%±16.0% vs. 43.1%±21.5%, $p=0.022$) in the H condition (Figure 6). The utilization of the remaining seven strategies did not differ between groups ($p>0.05$). Specifically in the H condition, HSA individuals changed strategies whereas the palpation strategy was used more in the SH condition compared to the H condition (38.8%±12.3% vs. 23.2%±16.1%, $p=0.013$). As sensory conditions changed, the time spent using strategies differed accordingly, in both the dynamic tracing (HSA: 0.5%±1.8% vs. 30.0%±22.8%, $p<0.001$; LSA: 0.0%±0.0% vs.
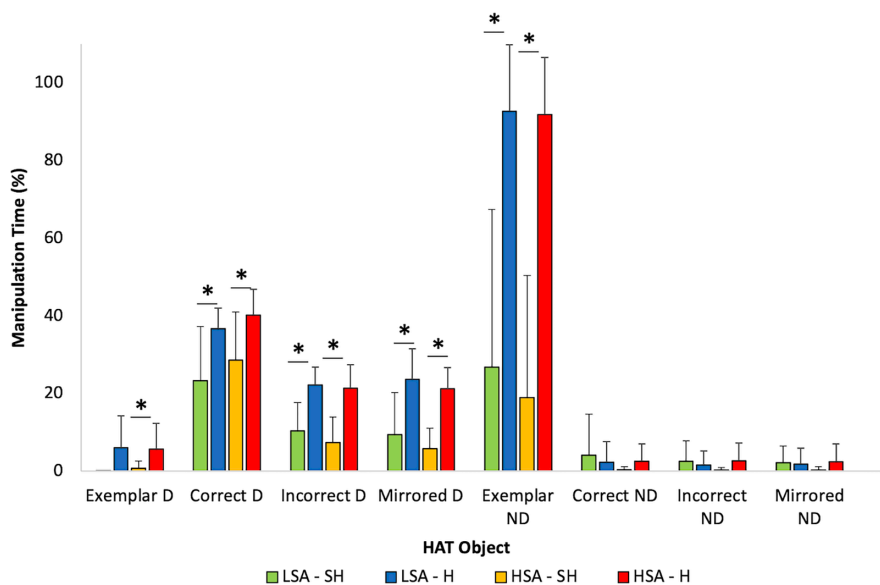

FIGURE 6 Percent of object manipulation time per question for each of the four HAT objects in the SH (pattern bars) and H condition (filled bars) using dominant (D) and non-dominant (ND) hands. No significant differences between the HSA and LSA ($n=21$) were found. However, across sensory modalities (SH-H) and within objects (Exemplar, Correct, Incorrect, Mirrored), objects were manipulated significantly less in the SH condition than in the H condition, by the dominant hand, in both HSA and LSA ($p<0.05$). The exemplar was manipulated significantly less in the dominant hand by the HSA group between the SH and H condition ($p<0.05$). The exemplar was manipulated significantly less in the non-dominant hand between the SH and H condition in both the HSA and LSA groups ($p<0.05$). Error bars indicate ±1*SD*. H, haptic; HAT, haptic abilities test; HSA, high spatial abilities, LSA, low spatial abilities; n, frequency; p, p-value; *SD*, standard deviation; SH, sighted haptic; %, percent. *$p<0.05$.

25.4% ± 17.1%, $p = 0.001$) and static hold (HSA: 0.0% ± 0.0% vs. 9.3% ± 9.4%, $p < 0.001$; LSA: 0.0% ± 0.0% vs. 9.3% ± 6.9%, $p = 0.002$) times were significantly lower in the SH condition compared to the H condition in both HSA and LSA individuals. Further, the lift strategy was used significantly less in the SH condition compared to

the H condition in the HSA group (8.7% ± 14.7% vs. 37.5% ± 36.1%, $p = 0.004$) (Figure 7).

DISCUSSION

Previous data from this laboratory suggest that the addition of haptic sensory information may aid LSA individuals in completing spatially and haptically challenging tasks.[22] Although test performance and timing of haptic behaviors may be similar between H/LSA individuals. The current inquiry regarding gaze and haptic behaviors may delineate how multimodal inputs (somatosensory and vision) contribute to problem-solving behaviors. This study systematically evaluated spatial and haptic abilities while deriving vision and action behavioral constructs. In line with the current hypothesis, both HSA and LSA individuals demonstrated similar gaze and hand behaviors when allowed to rely on haptic inputs in addition to their vision when completing a spatially challenging task.

Gaze behavior

In the previous study using the HAT construct, the speed of completion of spatially demanding tasks did not differ between HSA and LSA individuals.[22] The authors suggest that temporal differences alone do not provide sufficient detail to distinguish the variations in attentional behaviors that underlie individual strategies when confronted with spatially challenging tasks.[29] Pupillary fixations from the current study were similar between HSA and LSA groups, suggesting both sets of individuals rely on similar vision-based time
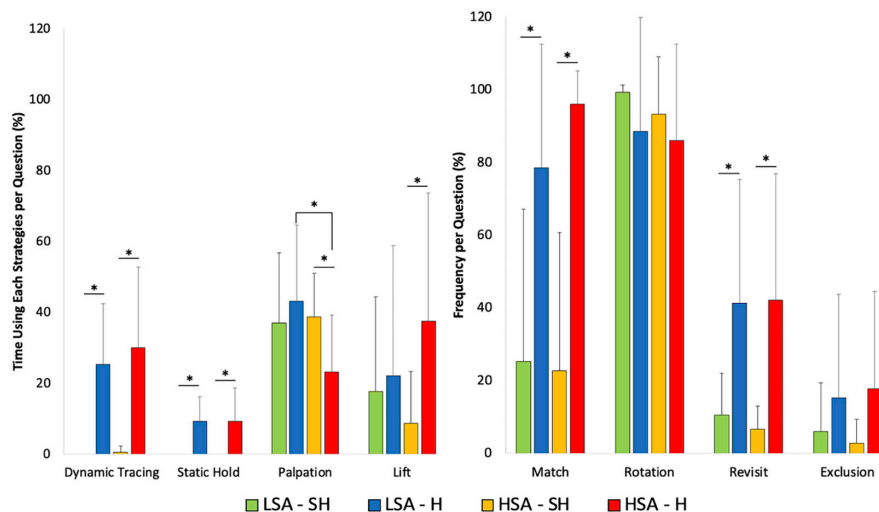
TABLE 2 Classification of haptic strategies.

Strategy	Definition
Dynamic tracing (%)	Object Tracing—discrete aspects of the object examined. Corners, lengths, and edges of the object actively explored while leaving the objects stationary
Static hold (%)	Hand Coverage—Hand placed over objects to apprehend fits entirety or may grab it without moving or lifting
Palpation (%)	Touching the object in a non-descript manner (i.e., moving the object, orienting it, knocking it over, and feeling to grasp it.)
Lift (%)	Elevating the object off the table with manipulation
Match (Y/N)	Putting two objects side by side for comparison
Rotation (Y/N)	Rotating the objects around the x, y, or z axis while keeping the object on the table or in the air
Revisiting (Y/N)	Feeling, touching and/or manipulating one block more than once per question
Exclusion (Y/N)	Moving a perceived incorrect match away from their comparators

Note: The % indicates the relative percentage of time spent depicting these behaviors per question while the binary Y/N indicates whether the behavior was expressed.

FIGURE 7 Haptic strategies employed during SH and H conditions of the HAT in HSA and LSA groups ($n = 22$). The proportion of time (average percent) of response time each strategy was employed per question. The palpation strategy was used more in the LSA group compared to the HSA groups ($p = 0.022$). There were no other significant differences between groups. The palpation strategy was used significantly more in the SH condition in the HSA group ($p = 0.013$). The dynamic tracing and static hold strategy were used significantly less in the SH condition in both groups ($p < 0.05$). The lift strategy was used significantly less in the SH condition in the HSA group ($p = 0.004$). H, haptic; HAT, haptic abilities test; HSA, high spatial abilities, LSA, low spatial abilities; n, frequency; p, p-value; SD, standard deviation; SH, sighted haptic; %, percent. *$p < 0.05$.

strategies when completing the HAT (Figure 3). Deeper analysis of gaze location however suggests that HSA and LSA individuals may "view" objects differently and we speculate that they attend to different salient locations on those objects during fixations (Figure 4). The resolution limitations of the head-free gaze tracking equipment employed in the current study did not enable intra-object examination as was undertaken previously in the series of studies by Roach et al.[29,56] In those studies, intra-object H/LSA gaze fixation locations differed on similar geometric blocks as those used in the current study. A difference between study methods is that Roach et al. used a computer screen to view the block shapes and a "head-fixed" gaze approach whereas here, the blocks were physical objects with head-free gaze tracking in the current study.

In the current study, LSA individuals spent more time fixating on the mirrored objects, which is an incorrect answer. These findings are indicative that LSA individuals devoted more cognitive processing to the mirrored match when answering the HAT questions in sighted conditions. This finding is in line with the hypothesis suggesting LSA students endure higher cognitive loads on identical learning and assessment exercises as their HSA counterparts.[57,58] The HSA groups recognized the mirrored match as incorrect, and did not afford further visual attention. This difference in visual attention between HSA and LSA groups appears to underlie differences in scores obtained in sighted conditions of the HAT with the HSA group scoring significantly higher than the LSA group. These findings may be compared to how students approach multiple-choice questions in written exams when using an elimination of distractors strategy.[59] Given scores were significantly lower in the LSA group in sighted conditions, and LSA individuals spent more time attending to the mirrored match, it is possible the LSA group's elimination strategy, if present, was less successful and lead the group to incorrectly identify the mirrored match as the correct answer. Similar findings were described in Roach et al.'s experiment, where temporal information combined with gaze location was used to indicate how H/LSA persons visually attend to different aspects within MRT images on a 2D computerized test.[29] Importantly, when information visual attention is afforded, the temporal information takes on deeper and more refined meanings regarding the salient aspects of attention.[29] Akin to present results, Roach et al. concluded that devoting equal temporal attention to locationally non-salient aspects will negatively affect performance, at least in terms of the visually derived MRT score.[29]

When haptic opportunities were afforded in the SH condition, differences in gaze location strategies disappeared between H/LSA groups and similar results were obtained on the HAT. These findings suggest underlying the mechanisms of performance degradations when haptics are removed from learning or assessment. Here, inclusion of haptics appears to elevate LSA individuals' performances on identical and spatially challenging tasks where attentional saliency may again play a role. In sighted conditions of this test and others,[29] LSA individuals sought for visual saliency, taking more time with a sporadic fixation approach compared to their HSA counterparts. Similarly, Van Polanen et al.[60] have conceptualized haptic saliency where features of objects may "pop out" with respect to other

object properties enabling efficient cognitive processing. Klatzky et al. suggest saliency extends beyond simple haptic properties of an object by describing qualities of immediate interactions with skin to global interactions of hand and object.[61] We suggest that the cumulative effects of haptic and visual saliency arrive at equivalent performance on the SH condition of the HAT, whereas performance differences still exist in the sighted only conditions, akin to those witnessed with performances on the mental rotations test.

Pupillometry

The pupillometry results from the current study also offer a unique vantage into underlying cognitive challenges that may be related to stress. Pupillary diameter is an indirect measure of an individual's stress.[62,63] Chen et al. suggest that changes in pupillary diameter are sensitive to cognitive load and communicative load, but they cannot index perceptual or physical loads.[63] In the current study, the percent change in pupillary diameter ($\Delta P\emptyset$) did not differ between HSA and LSA individuals in the SH or H condition (Figure 5) suggesting the addition of haptics does not induce stress responses differently in HSA or LSA individuals. If spatial ability is disregarded, the stress response was significantly lower in the H condition in both HSA and LSA individuals. Therefore, the addition of 3D objects affording haptic opportunities may explain equivalent stress responses in LSA individuals and may contribute to successfully completing spatially challenging tasks that are typically considered more challenging for LSA individuals. In contrast, in a study examining cerebrovascular responses during a learning and assessment exercise with 3D computer graphics, Loftus et al. observed small but significant drops in end-tidal carbon dioxide in LSA individuals. This was indicative of rising stress and hyperventilation, although mild, the consequences were declining brain blood flow in LSA, compared to HSA individuals undertaking the same 3D tests of anatomy.[64] Although not the main theme of that study, the authors suggest stress may have played a role in eliciting negative changes in LSA individuals more than HSA individuals through activation of the sympathetic nervous system resulting in mild ventilatory alterations that contribute to cerebrovascular vasoconstriction. That study's experimental testing method was driven entirely on a computer screen and did not involve haptics. Overall, it is striking that differing interactions of spatial and haptic abilities also interact with learning and/or testing modalities suggesting a contribution to unequal learner stressors that *appear unrelated* to the immediate task but represent an example of covert extraneous cognitive load[58] generated through an individual's perception of the visual environment and the objective of the tasks within said environment.

Hand behavior

The average proportion of manipulation time for each of the HAT objects between HSA and LSA individuals in both the SH and H

conditions was similar. Given that both HSA and LSA groups employed the same temporal strategies and obtained similar scores, three possible scenarios emerge; either HSA individuals' advantages were repressed, the addition of haptic input elevated LSA individual's abilities, or SA is not related to HA and were not required when completing the SH and H condition of the HAT.

However, there were clear differences in the cumulative proportion of time spent manipulating each of the HAT objects per question between the SH and H conditions. The time spent manipulating most HAT objects decreased in the SH condition compared to the H condition in both HSA and LSA. These findings suggest that when both visual and haptic information is afforded, subjects preferentially used visual sensory input over haptics to complete spatial tasks. This finding may be contextualized within the spatial properties of objects, including object shape and size, which are most efficiently recognized by one's sense of vision.[61] Furthermore, visual inputs tend to dominate the tactile interpretation of spatial properties of objects when there is a conflict between what is seen and what is felt.[65] Importantly, when visual information becomes less reliable, observers put progressively more decision-making emphasis on the information provided by haptics.[66] The combination of both visual and tactile information arising from one object promotes better integration of the object shape recognition.[66] Specifically, integrated visual–haptic inputs are more reliable than unimodal, visual or haptic, input.[66]

Although most haptic strategies employed during the SH and H conditions were similar in both SA groups, LSA individuals relied significantly more on the palpation strategy in the H condition. These findings support the notion that LSA individuals spend more time manipulating objects in a non-descript and haphazard fashion. This notion is not without precedent, in vision-based studies, some authors suggest LSA individuals who possess lower visuospatial working memory tend to use random exploration methods.[64,67] The random approach has consequences, more time is required to complete questions as increased cognitive loads are then encountered as LSA individuals must more frequently update fleeting mental representations to derive meaning.[64,67] Therefore, increased proportions of time spent utilizing the palpation strategy may reflect the requirement of more time for LSA individuals to correctly orient the object to problem solve.

In a prior study, vision-based spatial abilities measured with a redrawn Vandenberg and Kuse Mental Rotations Tests (MRT A&C) and a surface development test (SDT) was correlated with scores achieved on drawings created solely through blind haptic perception of 3D objects more complicated than those in the current study.[11] Moderate correlations have been found with MRT C and SDT. A weak correlation was found with MRT A, the test used in the current study. A statistically significant correlation between MRT and HAT score was not found in the study. Nonetheless, MRT scores were correlated to the palpation strategy during the HAT in the H condition. This may be an indication of a link between spatial abilities and haptic perception.

Performance in the haptic condition was similar between the HSA and LSA groups despite differing strategies between these groups. Together, the current findings indicate that the tactile inputs haptics may offer another sensory channel to offload or spread

sensory information and decreasing cognitive loads akin to the dual channels described in the multimedia learning theory.[1] Therein, the Dual Coding Theory (DCT) states that there are functional cognitive connections between verbal and non-verbal systems, specifically between verbal and pictorial inputs, and can help explain performance in intellectual tasks.[3,4] However, there are limitations to the DCT as it does not consider the possibilities of working memory support by different afferent information in the audiovisual inputs of words and images.[3,4] Therefore, the findings of this study illustrate how somatosensory inputs provided by haptics may represent a third sensory channel. We suggest that in scenarios where handling learning objects is part of the learning environment, like anatomy laboratories, there are multiple sensory coding channels available to the learner, sight, sound, and haptics. In this article, we begin to delineate how the addition of haptics may increase the reliability of object recognition when relevant sensory modalities (vision and haptics) are afforded. Importantly, haptic sensory channels appear to offer LSA individuals alternate sensory inputs through different strategies where object saliencies are presumably integrated into working memory. Research utilizing functional magnetic resonance imaging confirms a multichannel overlap between cortical areas subserving vision and haptics suggesting neural substrates are already in place to share object recognition.[48,68]

The current study elucidates whether, and how, haptic sensory inputs afford different input sources than those previously confined to vision and hearing in the Cognitive Theory of Multimedia Learning (CTML). Whether haptic sensory inputs append to the mental image models described in CTML[1] or exist cognitively as a haptic mental image within the working memory[7,69] is yet unknown.[70] The suggestion that LSA individuals in our paradigm successfully integrated haptically derived inputs appears plausible; although there were key differences in strategies regarding gaze location and haptic strategies, there were no differences in performance scores between HSA and LSA individuals during the SH and H conditions of the HAT. Within the construct of cognitive theory of multimedia learning, this study proposes haptics as a third sensory input. The current data suggest improvements in learning outcomes *are* achieved through one's sense of touch, at least with shape recognition. Common haptic strategies enable individuals to better select salient aspects of objects that contribute to accurate information transfer to working memory as mental representations, and presumably, long-term memory.

The current study expands the dual channel assumption of vision and audition in the CTML[4] to include important sensory information garnered through the hands as a haptic sensory input channel (Figure 8). The dual channel hypothesis states that humans possess separate information processing channels for these sensory inputs, and they are limited both in capacity[2] and active processing[1] capabilities where information is filtered, organized, schematized, and prepared for integration with long-term memories. The CTML states that if one sensory channel is overloaded, learning is reduced[5]; thus, spreading information over multiple sensory inputs is key to reduced cognitive overload. The current study cannot
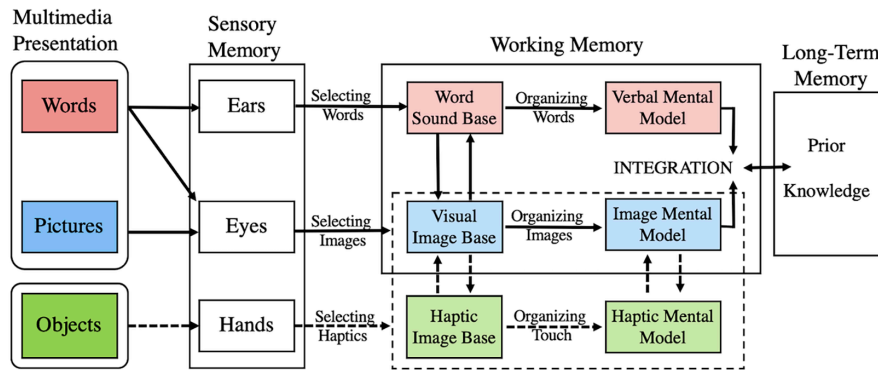
FIGURE 8 The Cognitive Theory of Multimedia Learning (CTML) words (red) and pictures (blue) are outlined in solid lines. In learning scenarios where touch is possible, the inclusion of haptic perception (green) provides viable sensory channel inputs for learning. The haptic input channel is appended to sensory input channels, when required, by the learner (outlined in broken line).

definitively address whether cognitive loads were different in the visual or haptic sensory conditions; however, we suggest that the addition of a third sensory input, haptics, would reduce cognitive load. The overload threat is reduced by decreased reliance on vision, and with it, the inherent need for spatial ability to succeed in these types of tasks. For those with high spatial ability (HSA), the need to touch may not be required since adequate salient information of object orientation is achieved through the visual channel.[29,58] For lower spatial ability (LSA) individuals, however, the use of haptics spares reliance on vision-based spatial abilities derived through the visual channel, enabling decisions to be made more easily. Given these results, care should be taken when diverting all sensory input to vision, even when better visualizations like stereoscopy and/or immersive environments are substituted in learning scenarios for non-expert learners.

Incorporating haptic opportunities in our teaching, learning, and assessment repertoires may raise LSA individual's performance by decreasing unnecessary cognitive load.[58] Given the requisite, and sometimes permanent, migrations to online anatomy teaching and learning environments due to Covid-19, a conundrum arises. Are educators unwittingly removing meaningful sensory fidelities from learners?[22,70,71] Gaining insight into learner behaviors when sensory inputs, like haptics, are eliminated provides valuable evidence that can guide the development of future anatomy curricula and educational resources.

Limitations and future directions

These findings are subject to limitations. The sample size did not enable comparisons between individuals with overtly divergent SA. Obtaining a larger sample size could allow for more divergent differences between H/LSA individuals as researchers could investigate whether greater divergence in performance, gaze behavior, and haptic behavior between groups having greater differences in baseline SA.

Further, the change in pupillometry diameter is variable, it is an indirect indicator of acute stress response possible due to

methodological concerns.[72–74] More direct measurements of stress may enable the identification of the underlying mechanisms contributing to our findings. Therefore, it may be worthwhile to better examine the stress response in individuals with diverging SA during the HAT to elucidate whether the addition of haptics elevates LSA individuals' performance through a decreased stress response.

As the HAT is an emerging paradigm, the current study did not include an a priori time constraint like the MRT, instead it was applied individually using a low-pass filtering approach. Here, an average time per question is calculated $\pm1SD$ for each HAT condition. This number represents an upper and lower time limit for each individual; if the time to answer is exceeded for a particular question, it is deemed incorrect. It is possible that an overt time constraint in the HAT, like that of the mental rotations test, would increase participants' stress and change the results found in HSA and LSA individuals. As the current study was exploratory in nature, the imposition of a standard time limit across all subjects may exacerbate hand–object behavioral differences, pupillometry, and HAT scores beyond what we witness here. This remains an area of future research opportunities as the field of haptic abilities grows.

Finally, to assess haptic behavior more discreetly during haptic tasks inherent to HAT, future studies may consider altering sensory feedback of the HAT shapes, recording more precise hand movements with haptic kinematic gloves, and/or time-motion video performance and data analysis software to better understand hand and or object motion tracking as it pertains to human perception meets learning behaviors. Further analysis of gaze location related to visual and haptic strategies used during the HAT may illustrate the further refinement of behavioral differences in individuals with divergent SA.

CONCLUSION

This study simultaneously evaluated gaze and hand behavior in spatially and haptically challenging tasks, based on the objects of the mental rotations test. Individual spatial abilities were inversely correlated to gaze fixation particularly in erroneous, but mirrored matches

in the S condition. This difference was eliminated when hands were used in the H condition, specifically the "palpation" strategy in the H condition of the HAT. The temporal measurements of gaze behavior of most indices (visual fixations per question, average fixation duration, and fixation rate) and haptic behavior (percentage of time spent touching each of the four HAT objects) were similar across spatial abilities indicating subtle and deeper differences in how spatial and haptic abilities afford environmental awareness. This is some of the first data to elucidate a third type of sensory memory to the information processing model contributing to the cognitive theory of multimedia learning, tactile sensory memory. Therefore, if opportunities for touch are included in learning and assessments scenarios, improved student performance could be a result.

AUTHOR CONTRIBUTIONS

Michelle A. Sveistrup: Formal analysis; investigation; methodology; project administration; writing – original draft; writing – review and editing. **Jean Langlois:** Investigation; supervision; visualization; writing – review and editing. **Timothy D. Wilson:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – original draft; writing – review and editing.

ACKNOWLEDGMENTS

ORCID

Jean Langlois https://orcid.org/0000-0002-8312-4273
Timothy D. Wilson https://orcid.org/0000-0001-7102-2491

REFERENCES

1. Mayer RE. Cognitive theory of multimedia learning. In: Mayer RE, editor. Cambridge handbook of multimedia learning. New York: Cambridge University Press; 2014. p. 31–48.
2. Sweller J, van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. Educ Psychol Rev. 1998;10(3):251–96.
3. Paivio A. Bilingual dual coding theory and memory. In: Heredia HR, Altarriba J, editors. Foundations of bilingual memory. New York: Springer New York; 2014. p. 41–62. https://doi.org/10.1007/978-1-4614-9218-4_3
4. Paivio A. Intelligence, dual coding theory, and the brain. Intelligence. 2014;47:141–58. https://doi.org/10.1016/j.intell.2014.09.002
5. Mayer RE, Moreno R. Nine ways to reduce cognitive load in multimedia learning. Educ Psychol. 2003 Mar 1;38(1):43–52.
6. Paas F, van Gog T, Sweller J. Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. Educ Psychol Rev. 2010;22(2):115–21.
7. Baddeley AD. Working memory. Philos Trans R Soc Lond Ser B Biol Sci. 1983;302:311–24.
8. Smith ER. Mental representation and memory. In: Gilbert DT, Fiske ST, Lindzey G, editors. The handbook of social psychology. Volume 1-2. 4th ed. New York: McGraw-Hill; 1998. p. 391–445.
9. James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA. Haptic study of three-dimensional objects activates extrastriate visual areas. Neuropsychologia. 2002;40(10):1706–14.
10. Jones LA, Lederman SJ. Prehension. In: Jones LA, Lederman SJ, editors. Human hand function. New York: Oxford University Press; 2006. p. 100–15. https://doi.org/10.1093/acprof:oso/9780195173154.003.0006
11. Langlois J, Hamstra SJ, Dagenais Y, Lecourtois M, Lemieux R, Yetisir E, et al. Objects drawn from haptic perception and vision-based spatial abilities. Anat Sci Educ. 2024;17:433–43.
12. Shepard RN, Metzler J. Mental rotation of 3D objects. Science. 1971;171:701–3.
13. McGee MG. Human spatial abilities: psychometric studies and environmental, genetic, hormonal, and neurological influences. Psychol Bull. 1979;86(5):889–918.
14. Uttal DH, Miller DI, Newcombe NS. Exploring and enhancing spatial thinking: links to achievement in science, technology, engineering, and mathematics? Curr Dir Psychol Sci. 2013;22(5):367–73.
15. Nguyen N, Nelson AJ, Wilson TD. Computer visualizations: factors that influence spatial anatomy comprehension. Anat Sci Educ. 2012;5(2):98–108.
16. Nguyen N, Mulla A, Nelson AJ, Wilson TD. Visuospatial anatomy comprehension: the role of spatial visualization ability and problem-solving strategies. Anat Sci Educ. 2014;7(4):280–8.
17. Rochford K. Spatial learning disabilities and underachievement among university anatomy students. Med Educ. 1985;19(1):13–26.
18. Roach VA, Mi M, Mussell J, Van Nuland SE, Lufler RS, DeVeau KM, et al. Correlating spatial ability with anatomy assessment performance: a meta-analysis. Anat Sci Educ. 2021;14(3):317–29.
19. Ekstrom RB, French JW, Harman HH, Dermen D. Manual for kit of factor-referenced cognitive tests. 1st ed. Princeton, NJ: Naval Research Contract: Educational Testing Service; 1976. p. 224.
20. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills. 1978;47(2):599–604.
21. Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C. A redrawn vandenberg and kuse mental rotations test - different versions and factors that affect performance. Brain Cogn. 1995;28:39–58.
22. Sveistrup MA, Langlois J, Wilson TD. Do our hands see what our eyes see? Investigating spatial and haptic abilities. Anat Sci Educ. 2023;16:756–67.
23. Carpenter R. Movements of the eyes. 2nd ed. London: Pion Ltd; 1988 593 p.
24. Galley N, Betz D, Biniossek C. Fixation durations—why are they so highly variable. 2015.
25. Dodge R. Five types of eye movements in the horizontal meridian plane of the field of regard. Am J Phys. 1903;8(4):307–29.
26. Aagten-Murphy D, Bays PM. Functions of memory across saccadic eye movements. Curr Top Behav Neurosci. 2019;41:155–83.
27. Hoang Duc A, Bays P, Husain M. Eye movements as a probe of attention. Prog Brain Res. 2008;171:403–11.
28. Just MA, Carpenter PA. Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. Cogn Psychol. 1985;92:137–72.
29. Roach VA, Fraser GM, Kryklywy JH, Mitchell DGV, Wilson TD. The eye of the beholder: can patterns in eye movement reveal aptitudes for spatial reasoning? Anat Sci Educ. 2015;9(4):357–66.
30. Lederman SJ, Klatzky RL. Tutorial review haptic perception: a tutorial. Atten Percept Psychophys. 2009;71(7):1439–59.
31. Klatzky RL, Lederman SJ, Metzger VA. Identifying objects by touch: "an expert system." Percept Psychophys. 1985;37(4):299–302.

32. Lederman SJ, Klatzky RL. Hand movements: a window in haptic object recognition. Cogn Psychol. 1987;19:342–68.

33. Kalisch T, Kattenstroth JC, Kowalewski R, Tegenthoff M, Dinse HR. Cognitive and tactile factors affecting human haptic performance in later life. PLoS One. 2012;7(1):e30420.

34. San Diego JP, Cox MJ, Quinn BFA, Newton JT, Banerjee A, Woolford M, et al. Researching haptics in higher education: the complexity of developing haptics virtual learning systems and evaluating its impact on students' learning. Comput Educ. 2012;59(1):156–66.

35. Williams RL, Chen M-Y, Seaton JM. Haptics-augmented simple-machine educational tools. J Sci Educ Technol. 2003;12(1):1–12.

36. Han I, Black JB. Incorporating haptic feedback in simulation for learning physics. Comput Educ. 2011;57(4):2281–90.

37. Skulmowski A, Rey GD. Measuring cognitive load in embodied learning settings. Front Psychol. 2017;8:1–6.

38. Hightower B, Lovato S, Davison J, Wartella E, Piper AM. Haptic explorers: supporting science journaling through mobile haptic feedback displays. Int J Hum Comput Stud. 2019;122(122):103–12.

39. Study NE. *The effectiveness of using the successive perception test I to measure visual-haptic tendencies in engineering students* [Doctoral Dissertation]. West Lafayette, IN: Purdue University; 2001.

40. Crandall R, Karadoğan E. Designing pedagogically effective haptic systems for learning: a review. Appl Sci. 2021;11(6245):1–29. https://doi.org/10.3390/app11146245

41. Zhou M, Jones DB, Schwaitzberg SD, Cao CGL. Role of haptic feedback and cognitive load in surgical skill acquisition. Proc Hum Factors Ergon Soc Annu Meet. 2007 Oct 1;51(11):631–5.

42. Kinnison T, Forrest ND, Frean SP, Baillie S. Teaching bovine abdominal anatomy: use of a haptic simulator. Anat Sci Educ. 2009;2(6):280–5.

43. Jones MG, Minogue J, Tretter TR, Negishi A, Taylor R. Haptic augmentation of science instruction: does touch matter? Sci Educ. 2006;90(1):111–23.

44. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities and anatomy knowledge assessment: a systematic review. Anat Sci Educ. 2017;10(3):235–41.

45. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities and technical skills performance in health care: a systematic review. Med Educ. 2015;49:1065–85.

46. Antonoff MB, Swanson JA, Green CA, Mann BD, Maddaus MA, D'Cunha J. The significant impact of a competency-based preparatory course for senior medical students entering surgical residency. Acad Med. 2012;87(3):308–19.

47. Louridas M, Szasz P, De Montbrun S, Harris KA, Grantcharov TP. Can we predict technical aptitude?: a syste matic review. Ann Surg. 2016;263(4):673–91.

48. Snow JC, Strother L, Humphreys GW. Haptic shape processing in visual cortex. J Cogn Neurosci. 2014;26:1154–67.

49. Chow JK, Palmeri TJ, Gauthier I. Haptic object recognition based on shape relates to visual object recognition ability. Psychol Res. 2022;86:1262–73.

50. Tiffin J, Asher EJ. The Purdue Pegboard: norms and studies of reliability and validity. J Appl Psychol. 1948;32(3):234–47.

51. Toth AJ, Campbell MJ. Investigating sex differences, cognitive effort, strategy, and performance on a computerised version of the mental rotations test via eye tracking. Sci Rep. 2019;9(1):1–11.

52. Charmaz K. Constructing grounded theory: a practical guide through quantitative analysis. London: Sage Publications Ltd; 2006.

53. Creswell JW. Research design: qualitative, quantitative and mixed methods approaches. 3rd ed. Thousand Oaks, CA: SAGE Publications Inc; 2009. p. 224.

54. Elliot JM, Connolly KJ. A classification of manipulative hand movements. Dev Med Child Neurol. 1984;26:283–96.

55. Jones LA, Lederman SJ. Active haptic sensing. Human hand function. online ed. New York: Oxford University Press; 2006. p. 75–99.

56. Roach VA, Fraser GM, Kryklywy JH, Mitchell DGV, Wilson TD. Time limits in testing: an analysis of eye movements and visual attention in spatial problem solving. Anat Sci Educ. 2017;10(6):528–37.

57. Huk T. Who benefits from learning with 3D models? The case of spatial ability. J Comput Assist Learn. 2006;22:392–404.

58. Wilson TD. The role of image and cognitive load in anatomical multimedia. In: Chan LK, Pawlina W, editors. Teaching anatomy: a practical guide. 2nd ed. Cham: Springer Nature; 2020. p. 301–11.

59. Katalayi GB. Elimination of distractors: a construct-irrelevant strategy? An investigation of examinees' response decision processes in an EFL multiple-choice reading test. Theory Pract Lang Stud. 2018;8(7):749.

60. Van Polanen V, Bergmann Tiest WM, Kappers AML. Target contact and exploration strategies in haptic search. Sci Rep. 2014;4:1–7.

61. Klatzky RL, Lederman SJ, Reed C. There's more to touch than meets the eye: the salience of object characteristics with and without vision. J Exp Psychol Gen. 1987;116(4):299–302.

62. Polt JM. Effect of threat of shock on pupillary response in a problem-solving situation. Percept Mot Skills. 1970;31(2):587–93.

63. Chen S, Epps J, Paas F. Pupillometric and blink measures of diverse task loads: implications for working memory models. Br J Educ Psychol. 2023;93 Suppl 2:318–38.

64. Loftus JJ, Jacobsen M, Wilson TD. A view of cognitive load through the lens of cerebral blood flow. Br J Educ Technol. 2017;48:1030–46.

65. Rock I, Victor J. Vision and touch: an experimentally created conflict between the two senses. Science. 1964;143(3606):594–6.

66. Helbig HB, Ernst MO. Optimal integration of shape information from vision and touch. Exp Brain Res. 2007;179(4):595–606.

67. Shah P, Freedman EG. Visuospatial cognition in electronic learning. J Educ Comput Res. 2003 Oct 1;29(3):315–24.

68. Sathian K. Analysis of haptic information in the cerebral cortex. J Neurophysiol. 2016;116(4):1795–806.

69. Baddeley A. Working memory. Science. 1992;255(5044):556–9.

70. Wilson TD. Visualisation technologies-I can see clearly now but the feel is gone: Commentary on: Stereoscopic three-dimensional visualisation technology in anatomy learning: A meta-analysis, Bogomolova et al. Med Educ. 2021;55(33):285–8.

71. Evans DJR, Bay BH, Wilson TD, Smith CF, Lachman N, Pawlina W. Going virtual to support anatomy education: a STOPGAP in the midst of the Covid-19 pandemic. Anat Sci Educ. 2020;13(3):279–83.

72. Lee E-H. Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res (Korean Soc Nurs Sci). 2012;6(4):121–7.

73. Pedrotti M, Mirzaei MA, Tedescho A, Chardonnet J-R, Merienne F. Automatic stress classification with pupil diameter analysis. Int J Hum Comput Interact. 2014;30(3):220–36.

74. Uba J, Nuamah J. Investigating human physiological responses to work-related stress. Proc Hum Factors Ergon Soc Annu Meet. 2023;67(1):2285–90. Washington, DC. https://doi.org/10.1177/21695067231192596

AUTHOR BIOGRAPHIES

Michelle A. Sveistrup, M.Sc. (Clinical Anatomy), is now a second-year medical student at McGill University, Montreal, Quebec, Canada. She was an active member of the CRIPT Laboratory at the University of Western Ontario, London, Ontario, Canada. She investigated the relationships between spatial and haptic abilities and their associated attentional behaviors.

Jean Langlois, M.D., M.Sc. (Anatomy), C.S.P.Q., F.R.C.P.C. (Emergency Medicine), is an active member in the Department of Emergency Medicine, CIUSSS de l'Estrie – Centre hospitalier universitaire de Sherbrooke, Sherbrooke, Quebec, Canada. His field of research is the relationships between spatial abilities, anatomy knowledge, and technical skills performance in health care. He is also investigating the relationship between spatial and haptic abilities.

Timothy D. Wilson, Ph.D., is an associate professor in the Departments of Anatomy and Cell Biology and Dentistry at the University of Western Ontario, London, Ontario, Canada. He is the director of the gross anatomical curriculum for dentistry and directs the Corps for Research of Instructional and Perceptual Technologies (CRIPT) Laboratory. He studies the neurophysiological mechanisms of haptics and vision and consequent behaviors underlying learning.