

Response of Thermocouples Interfaced to Electrothermometers When Immersed in 5 Water Bath Temperatures

Blaine C. Long, PhD, LAT, ATC*; Lisa S. Jutte, PhD, LAT, ATC†; Kenneth L. Knight, PhD, ATC, FNATA, FACSM‡

*Department of Health and Human Performance, Oklahoma State University, Stillwater; †School of Physical Education, Sport, & Exercise Science, Ball State University, Muncie, IN; ‡Human Performance Research Center, Brigham Young University, Provo, UT

Context: Thermocouples and electrothermometers are used in therapeutic modality research. Until recently, researchers assumed that these instruments were valid and reliable.

Objective: To examine 3 different thermocouple types in 5°C, 15°C, 18.4°C, 25°C, and 35°C water baths.

Design: Randomized controlled trial.

Setting: Therapeutic modality laboratory.

Intervention(s): Eighteen thermocouple leads were inserted through the wall of a foamed polystyrene cooler. The cooler was filled with water. Six thermocouples (2 of each model) were plugged into the 6 channels of the Datalogger and 6 randomly selected channels in the 2 Iso-Thermexes. A mercury thermometer was immersed into the water and was read every 10 seconds for 4 minutes during each of 6 trials. The entire process was repeated for each of 5 water bath temperatures (5°C, 15°C, 18.4°C, 25°C, 35°C).

Main Outcome Measure(s): Temperature and absolute temperature differences among 3 thermocouple types (IT-21, IT-18, PT-6) and 3 electrothermometers (Datalogger, Iso-

Thermex calibrated from –50°C to 50°C, Iso-Thermex calibrated from –20°C to 80°C).

Results: Validity and reliability were dependent on thermocouple type, electrothermometer, and water bath temperature ($P < .001$; modified Levene $P < .05$). Statistically, the IT-18 and PT-6 thermocouples were not reliable in each electrothermometer; however, these differences were not practically different from each other. The PT-6 thermocouples were more valid than the IT-18s, and both thermocouple types were more valid than the IT-21s, regardless of water bath temperature ($P < .001$).

Conclusions: The validity and reliability of thermocouples interfaced to an electrothermometer under experimental conditions should be tested before data collection. We also recommend that investigators report the validity, the reliability, and the calculated uncertainty (validity + reliability) of their temperature measurements for therapeutic modalities research. With this information, investigators and clinicians will be better able to interpret and compare results and conclusions.

Key Words: reliability, validity, uncertainty

Key Points

- The validity, reliability, and uncertainty of temperature measurement depended on the thermocouple and electrothermometer type.
- Extreme water bath temperatures decreased the validity, reliability, and uncertainty of temperature measurement.
- Small-diameter thermocouples were more uncertain and less valid and reliable than large-diameter thermocouples, especially at extreme water bath temperatures.
- Thermocouples interfaced with Iso-Thermex electrothermometers provided more- valid, more-reliable, and lower uncertainty values than those interfaced with the Datalogger electrothermometer.
- Researchers should measure and report the validity, reliability, and calculated uncertainty of temperature measurements so readers can better compare studies.

Thermocouples and electrothermometers are instruments used to measure tissue temperature during therapeutic modality research. Until recently, investigators assumed that these instruments were valid and reliable.^{1,2} Valid instruments represent the extent to which data are correct or true,³ whereas reliable instruments yield the same results on repeated trials.⁴ During therapeutic modality research, using an invalid or unreliable thermocouple or electrothermometer could lead to false interpretations of the data and influence the results or conclusions.

In 2 previous investigations,^{1,2} the authors suggested that investigators examining tissue temperature should

know the manufacturers' claims of uncertainty, test their equipment, and report validity and reliability of their instruments. However, these suggestions were based on a single room-temperature water bath rather than multiple temperatures. In addition, this single water bath temperature did not represent an important temperature for therapeutic modalities research, such as intramuscular or modality temperatures. Therefore, the purpose of our study was to determine if the validity, reliability, and uncertainty of 3 different thermocouple types vary among the 5 water bath temperatures.

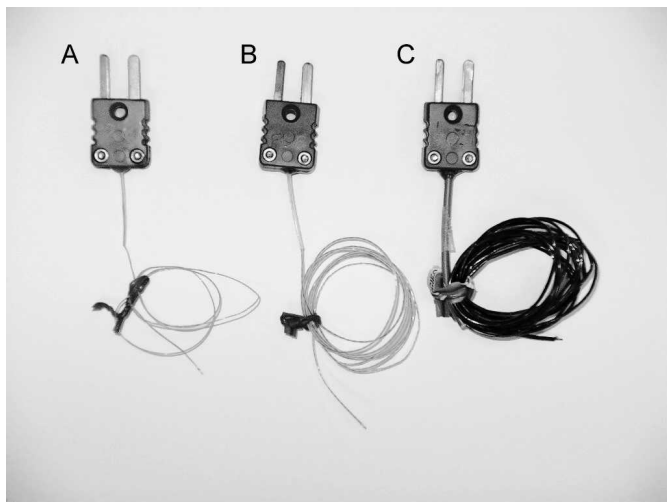


Figure. The 3 thermocouple types tested within the 3 electrothermometers (Physitemp Instruments Inc, Clifton, NJ). A, Implantable IT-21. B, Implantable IT-18. C, Surface PT-6.

METHODS

Study Design

A $3 \times 4 \times 5 \times 25$ randomized controlled laboratory experiment guided data collection. The independent variables were thermocouple type (PT-6, IT-18, IT-21 [Physitemp Instruments Inc, Clifton, NJ]), thermometer (mercury thermometer, Iso-Thermex –50:50, Iso-Thermex –20:80 [Columbus Instruments, Columbus, OH]), Datalogger [model MMS 3000-T6V4; Commtest Instruments Ltd, Christchurch, NZ]), temperature (5°C , 15°C , 18.4°C , 25°C , 35°C), and time (every 10 seconds for 4 minutes). The dependent variables were the absolute difference between thermocouple and mercury thermometer measurements (validity) and the SD of the individual thermocouple temperature measurements (reliability).

Instruments

Eighteen thermocouples (6 each of the 3 different thermocouple types: PT-6, IT-21, and IT-18) were examined (Figure). For each thermocouple type, 2 thermocouples were new and had not been used in previous studies, and 4 thermocouples had been used in previous studies. Thermocouples were interfaced to (1) a 6-channel Datalogger with a temperature range of -250°C to 350°C , (2) a 16-channel Iso-Thermex with a temperature range of -50°C to 50°C (Iso –50:50), and (3) a 16-channel Iso-Thermex with a temperature range of -20°C to 80°C (Iso –20:80). A mercury thermometer (model 15-059-18; Fisher Scientific, Pittsburgh, PA) certified by the National Institute of Standards and Technology (NIST) and graded at 0.1°C was used to measure water bath temperature. A Corning stirrer (model PC 103; Corning Inc, Corning, NY) and magnetic stir bar circulated the water bath.

Procedures

Eighteen thermocouple leads in 3 rows of 6 were inserted through the wall of a $23 \times 15 \times 19$ -cm foam polystyrene cooler and secured with silicone polymer. The thermocou-

ples extended approximately 10 cm into the cooler. The bottom row was 4 cm from the bottom of the cooler. Thermocouples within a row and between rows were 3 cm apart. The cooler was filled with water to within 3 cm of its top. The mercury thermometer was immersed into the water to approximately 3 cm from the bottom of the cooler. The thermocouples and mercury thermometer remained in place for the duration of the study.

Six thermocouples (2 of each model) were plugged into the 6 channels of the Datalogger and 6 randomly selected channels in the 2 Iso-Thermexes. Data were collected simultaneously from the 3 electrothermometers, whereas the mercury thermometer was read by the same investigator (L.S.J.) every 10 seconds for 4 minutes during each of 6 trials. The entire process was repeated for each of the 5 water bath temperatures (5°C , 15°C , 18.4°C , 25°C , 35°C). The temperature ranges were selected to represent temperatures commonly observed in therapeutic modality research (ie, 5°C for cryotherapy, 18.4°C for room temperature, and 35°C for normal intramuscular temperature). To ensure that water bath temperatures remained stable during data collection, water volume was held constant, and water temperature was monitored every second with the NIST-certified mercury thermometer. When necessary, water bath temperature was adjusted by adding heated water or crushed ice between trials.

Statistical Analysis

We calculated the means and SDs of the 3 thermocouple types, 3 electrothermometers, and 5 water bath temperatures, and we calculated the absolute difference between thermocouple and mercury thermometer measurements. The mean difference between the mercury thermometer and thermocouple measurements was our measure of validity.^{1,2} Differences in validity were analyzed using 3-way, repeated-measures analyses of variance (ANOVAs) for electrothermometer, water temperature, and thermocouple type. Next, we used 3 2-way, repeated-measures ANOVAs for thermocouple type and water temperature. Scheffé multiple comparison tests identified differences within our interaction term.

The SD of thermocouple measurements served as the measure of reliability.^{1,2} Differences in reliability among thermocouple, electrothermometer, and water bath temperatures were analyzed with modified Levene equal variance tests⁵ and pairwise *F* tests when appropriate.⁵

Uncertainty was calculated as the sum of validity and reliability.^{1,2} Differences less than the thermocouple manufacturer's reported claim ($\pm 0.1^{\circ}\text{C}$) were considered not practically different.

Means, SDs, and reliability calculations were completed with Microsoft Excel 2003 (Microsoft Corporation, Redmond, WA). We used Number Crunchers Statistical Software (NCSS 2001; NCSS, Kaysville, UT) for all analytical statistics. The α level was set at .05.

RESULTS

Thermocouple validity and reliability were dependent on thermocouple type and water bath temperature ($F_{16,36} = 17.03$, $P < .001$; modified Levene $P < .05$). Thermocouples were more valid in the Iso-Thermex units than in the Datalogger (Scheffé $P < .05$). Regardless of electrother-

Table 1. Validity and Reliability of Thermocouple Types Across Different Water Bath Temperatures for the Iso -50:50^a Electrothermometer (n = 162 Measurements)

	Temperature, °C	Thermocouple Type ^b		
		PT-6	IT-18	IT-21
Absolute thermocouple-mercury thermometer differences, mean (validity) ± SD	5	0.13 ^c ± 0.02	0.14 ^c ± 0.10	0.17 ^c ± 0.07
	15	0.12 ^c ± 0.02	0.14 ^c ± 0.01	0.13 ^c ± 0.09
	18.4	0.12 ^c ± 0.04	0.16 ^c ± 0.02	0.14 ^c ± 0.09
	25	0.05 ^c ± 0.05	0.08 ^c ± 0.02	0.33 ^c ± 0.22
	35	0.06 ^c ± 0.04	0.09 ^c ± 0.03	0.20 ^c ± 0.16
Thermocouple measurement reliability, mean ± SD	5	4.87 ± 0.02	4.84 ± 0.10	4.94 ± 0.17 ^d
	15	14.88 ± 0.02	14.85 ± 0.01	14.92 ± 0.13
	18.4	18.27 ± 0.04	18.24 ± 0.02	18.24 ± 0.14
	25	24.95 ± 0.05	24.92 ± 0.02	25.20 ± 0.34 ^{d,e}
	35	34.94 ± 0.04	34.90 ± 0.03	35.01 ± 0.25 ^{d,e}

^a Iso-Thermex calibrated from -50°C to 50°C (Columbus Instruments, Columbus, OH).

^b Physitemp Instruments Inc, Clifton, NJ.

^c Indicates PT-6 thermocouples were more valid than IT-18 thermocouples, and both PT-6 and IT-18 thermocouples were more valid than IT-21 thermocouples.

^d Indicates PT-6 thermocouples were more reliable than IT-21 thermocouples at 5°C, 25°C, and 35°C.

^e Indicates IT-18 thermocouples were more reliable than IT-21 thermocouples at 25°C and 35°C.

meter, the IT-18 and PT-6 thermocouples were not reliable; however, these differences were less than 0.1°C, which were the smallest differences we could detect, so they were not meaningful.

The validity and reliability for the Iso -50:50 are summarized in Table 1. Regardless of water bath temperature, the PT-6 thermocouples were more valid than the IT-18s, and both thermocouple types were more valid than the IT-21s ($F_{2,3} = 128.45, P < .001$). The IT-21 thermocouples were not as reliable as the PT-6 thermocouples at 5°C ($F_{53,53} = 92.33, P < .001$), 25°C ($F_{53,53} = 19.66, P < .001$), and 35°C ($F_{53,53} = 85.44, P < .001$) and were not as reliable as the IT-18s at 25°C ($F_{53,53} = 292.41, P < .001$) and 35°C ($F_{53,53} = 87.80, P < .001$).

The validity and reliability for the Iso -20:80 are summarized in Table 2. Validity differed across thermocouple type and water bath temperature in the Iso -20:80 ($F_{8,12} = 7.31, P < .001$). The IT-21 thermocouples were less valid at 35°C than at 15°C and 18.4°C (Scheffé $P < .05$). The IT-21 thermocouples at 35°C were also less

valid than the PT-6 thermocouples at 15°C, 18.4°C, 25°C, and 35°C and the IT-18 thermocouples at 15°C, 18.4°C, and 25°C (Scheffé $P < .05$). The IT-21 thermocouples at 25°C and 35°C were not as reliable as the PT-6 thermocouples at 25°C ($F_{53,53} = 78.03, P < .001$) and 35°C ($F_{53,53} = 195.53, P < .001$), respectively. In addition, the IT-21 thermocouples at 25°C and 35°C were not as reliable as the IT-18s at 25°C ($F_{53,53} = 246.61, P < .001$) and 35°C ($F_{53,53} = 756.25, P < .001$).

The validity and reliability for the Datalogger are summarized in Table 3. We found an interaction between thermocouple type and water bath temperature in the Datalogger ($F_{8,12} = 3.62, P = .02$). The IT-21 thermocouples at 35°C were less valid than the PT-6 and IT-21s at 18.4°C (Scheffé $P < .05$). The IT-21s at 25°C and 35°C were less valid than the IT-18s at 15°C and 18.4°C (Scheffé $P < .05$). The IT-21 thermocouples were not as reliable as PT-6 thermocouples at 25°C ($F_{53,53} = 80.47, P < .001$) and IT-18 thermocouples at 25°C ($F_{53,53} = 28.17, P < .001$) and 35°C ($F_{53,53} = 21.66, P < .001$).

Table 2. Validity and Reliability of Thermocouple Types Across Different Water Bath Temperatures for the Iso -20:80^a Electrothermometer (n = 162 Measurements)

	Temperature, °C	Thermocouple Type ^b		
		PT-6	IT-18	IT-21
Absolute thermocouple-mercury thermometer differences, mean (validity) ± SD	5	0.18 ± 0.02	0.16 ± 0.02	0.15 ± 0.05
	15	0.08 ^c ± 0.02	0.11 ^c ± 0.02	0.03 ± 0.02 ^d
	18.4	0.05 ^c ± 0.02	0.12 ^c ± 0.02	0.02 ± 0.01 ^d
	25	0.07 ^c ± 0.04	0.03 ^c ± 0.02	0.71 ± 0.42
	35	0.11 ^c ± 0.06	0.02 ± 0.02	1.0 ^{c,d} ± 0.82
Thermocouple measurement reliability, mean ± SD	5	4.82 ± 0.02	4.84 ± 0.02	5.04 ± 0.15
	15	14.92 ± 0.02	14.89 ± 0.02	14.98 ± 0.03
	18.4	18.35 ± 0.03	18.28 ± 0.02	18.39 ± 0.02
	25	25.06 ± 0.05 ^e	24.98 ± 0.03 ^e	25.71 ± 0.42 ^e
	35	35.11 ± 0.06 ^e	34.99 ± 0.03 ^e	36.02 ± 0.83 ^e

^a Iso-Thermex calibrated from -20°C to 80°C (Columbus Instruments, Columbus, OH).

^b Physitemp Instruments Inc, Clifton, NJ.

^c Indicates IT-21 thermocouples were less valid at 35°C than PT-6 thermocouples at 15°C, 18.4°C, 25°C, and 35°C and IT-18 thermocouples at 15°C, 18.4°C, and 25°C.

^d Indicates IT-21 thermocouples were less valid at 35°C than at 15°C and 18.4°C.

^e Indicates IT-21 thermocouples were less reliable at 25°C and 35°C than PT-6 thermocouples at 25°C and 35°C and IT-18 thermocouples at 25°C and 35°C.

Table 3. Validity and Reliability of Thermocouple Types Across Water Bath Temperatures for the Datalogger^a (n = 162 Measurements)

	Temperature, °C	Thermocouple Type ^b		
		PT-6	IT-18	IT-21
Absolute thermocouple–mercury thermometer differences, mean (validity) ± SD	5	0.93 ± 0.04	0.87 ± 0.05	0.96 ± 0.16
	15	0.75 ± 0.05	0.69 ^c ± 0.03	0.82 ± 0.18
	18.4	0.71 ^d ± 0.03	0.65 ^{c,d} ± 0.05	0.67 ± 0.10
	25	0.79 ± 0.02	0.79 ± 0.04	1.00 ^c ± 0.20
	35	0.80 ± 0.15	0.80 ± 0.04	1.00 ^{c,d} ± 0.20
Thermocouple measurement reliability, mean ± SD	5	5.93 ± 0.04	5.87 ± 0.05	5.95 ± 0.16
	15	15.75 ± 0.05	15.69 ± 0.03	15.81 ± 0.18
	18.4	19.11 ± 0.03	19.05 ± 0.05	19.07 ± 0.10
	25	25.79 ± 0.02 ^e	25.79 ± 0.04 ^e	26.00 ± 0.21 ^e
	35	35.79 ± 0.15	35.80 ± 0.04 ^e	36.00 ± 0.20

^a Commtest Instruments Ltd, Christchurch, NZ.

^b Physitemp Instruments Inc, Clifton, NJ.

^c Indicates IT-21 thermocouples were less valid at 25°C and 35°C than IT-18 thermocouples at 15°C and 18.4°C.

^d Indicates IT-21 thermocouples were less valid at 35°C than PT-6 and IT-21 thermocouples at 18.4°C.

^e Indicates IT-21 thermocouples were less reliable than PT-6 thermocouples at 25°C and IT-18 thermocouples at 25°C and 35°C.

Based on our validity and reliability, we calculated the uncertainty (validity + reliability) for each thermocouple type in each electrothermometer at the 5 water baths. These values are presented in Table 4.

DISCUSSION

The validity and reliability of temperature measurements can be influenced by factors separate from the thermocouple type¹ and electrothermometer.² Our extreme water bath temperatures (5°C and 35°C) were more likely than other water bath temperatures to decrease the validity and reliability of temperature measurement, especially with the smaller IT-21 thermocouples. In addition, our observations support those of previous investigations in which temperature measurement validity and reliability depended on the equipment used.^{1,2} The PT-6 and IT-18 thermocouples used with an Iso –50:50 and Iso –20:80 had similar validity and reliability,¹ whereas the PT-6 interfaced with a Datalogger was less valid, was less reliable, and had a higher uncertainty value.²

Table 4. Uncertainty for Thermocouple Types and Water Bath Temperature Within Each Electrothermometer

	Temperature, °C	PT-6	IT-18	IT-21
Iso –50:50 ^a	5	±0.15	±0.24	±0.34
	15	±0.14	±0.15	±0.26
	18.4	±0.16	±0.18	±0.28
	25	±0.10	±0.10	±0.67
	35	±0.10	±0.12	±0.45
Iso –20:80 ^b	5	±0.20	±0.18	±0.30
	15	±0.10	±0.13	±0.06
	18.4	±0.08	±0.14	±0.04
	25	±0.12	±0.06	±1.13
	35	±0.17	±0.05	±1.83
Datalogger ^c	5	±0.97	±0.92	±1.12
	15	±0.80	±0.72	±1.00
	18.4	±0.74	±0.70	±0.77
	25	±0.81	±0.83	±1.21
	35	±0.95	±0.84	±1.20

^a Iso-Thermex (Columbus Instruments, Columbus, OH) calibrated from –50°C to 50°C.

^b Iso-Thermex calibrated from –20°C to 80°C.

^c Commtest Instruments Ltd, Christchurch, NZ.

Within each electrothermometer, the major differences in temperature validity and reliability were observed with the IT-21 thermocouples at extreme temperatures. In the Iso-50:50, the IT-21 thermocouples were not as valid as the PT-6 thermocouples in the coldest water bath (5°C), and they were not as valid at the warmest water baths (25°C and 35°C).

Although they may be more uncertain than other types of thermocouples, the IT-21 thermocouples often are used to measure the most pertinent temperature, which is intramuscular temperature. Differences in validity are likely due to the small-diameter wire of the IT-21s. From our observations, investigators often assume that the IT-21, which is implanted with a 21-gauge hypodermic needle, produces less pain during insertion than the large-diameter implantable thermocouple types (ie, multi-sensor probes or IT-18s), which must be implanted with a larger-gauge needle or catheter. However, this small diameter also decreases the durability of the IT-21s, making them more fragile and susceptible to unnoticed damage. Unnoticed damage from previous use may account for the decreases in validity and reliability for some of our IT-21 thermocouples.

The PT-6 thermocouples interfaced to the Iso –50:50 were more valid than the IT-18 thermocouples. This difference was likely due to the large SD (±0.10) observed with the IT-18 thermocouples in the 5°C water bath (Table 1). This large SD may have occurred because IT-18 thermocouples are designed for in vivo measurements, which average approximately 33°C. During cryotherapy research, intramuscular temperatures only decrease about 7°C^{6–8} to 10°C⁹ when a crushed ice bag is applied directly to the skin. Even these in vivo temperatures are much greater than 5°C. It appears that the IT-18 thermocouples were not as valid at lower temperatures, such as 5°C, and that IT-18 thermocouples are more appropriate for measuring in vivo temperatures rather than skin surface temperatures during cryotherapy treatments.

We cannot compare our validity or reliability data with manufacturers' claims because the electrothermometer manufacturers report only accuracy,^{10,11} and the thermocouple manufacturers report only the maximal temperature rather than validity or reliability. The maximal temperature

for both the IT-21 and IT-18 is 150°C and for the PT-6 is 400°C.¹²

Although the use of an analog mercury thermometer could have introduced human error, we attempted to minimize this error and maintain that human error would be less than 0.1°C. First, we used an NIST-certified thermometer, which was graded to 0.1°C. The NIST certification provides the scientist with a level of confidence regarding the mercury thermometer. Second, we structured our methods to minimize error not associated with the instrument itself. Therefore, the same investigator read the mercury thermometer for all measurements, thus eliminating interrater error. Any human reading error would be less than 0.1°C because the thermometer was graded to 0.1°C. Third, we did not attempt to evaluate validity or reliability more precisely than the capability of our analog thermometer or possible amount of human error (0.1°C).

Consumers of therapeutic modality research should be cautious regarding those studies in which the specific validity and reliability for the temperature measurements are not reported. Depending on the instrumentation used and temperatures measured, an additional 0.2°C to 1.0°C of variation can be added to reported temperatures if one applies our uncertainty values. In the best of circumstances, there would be little effect on the meaning of the data. For example, skin temperature reported by Merrick et al¹³ using a TX-31 (Columbus Instruments), which is comparable to a PT-6, indicated a 0.16°C difference between an ice bag and Flex-i-cold (Cramer Products Inc, Gardner, KS) treatments. Standard deviations were $\pm 1.36^\circ\text{C}$ and $\pm 1.56^\circ\text{C}$, respectively. Based on our uncertainty data, the greatest difference between treatments could be 0.36°C (Table 4). This would have minimal effect on the interpretation of skin temperature in this experiment, whereas in other circumstances, it could have a major effect on the data reported. When researchers compared 3 therapeutic ultrasound units (ie, Omnisound 3000C [Accelerated Care Plus Corporation, Reno, NV], Dynatron 950 [Dynatronics Corporation, Salt Lake City, UT], and Excel Ultra III [XLTEK, Oakville, ON]), they could not increase intramuscular tissue temperature above the therapeutic threshold of 40°C with the Dynatron 950 or Excel Ultra III.¹⁴ Applying our uncertainty calculation for the most similar data collection condition (Table 4), a Datalogger with an IT-21 (most similar to the IT-23 used) at 35°C would add $\pm 1.2^\circ\text{C}$ of uncertainty. If the final tissue temperature in those treated with the Dynatron 950 or Excel Ultra III units increased by more than 1°C, these groups would have reached the therapeutic threshold for ultrasound and, therefore, altered part of the conclusion based on these data.

We make the following recommendations for researchers who want to collect temperature data. First, they should use an Iso-Thermex electrothermometer instead of a Datalogger unit. Second, because large-diameter thermocouples (ie, IT-18s) are less uncertain than small-diameter thermocouples (ie, IT-21s), researchers should use the largest implantable thermocouple possible for intramuscular data collection. Keep in mind that the larger thermocouples also have larger time constants, meaning more time is needed to determine the step change in the temperature measure.¹² All of the thermocouple types used in this study had a time constant of less than 0.1 second;

therefore, temperature was determined in 0.5 seconds. Third, because the specific circumstance surrounding temperature collection can influence validity and reliability, researchers should report the validity and reliability in their test conditions so readers can better compare studies.

Most authors of reliability studies have reported intraclass correlation coefficients (ICCs) for their data. This would be inappropriate for our data because testing was conducted on several homogenous populations (ie, stable water baths). Intraclass correlation coefficients are used to assess the variability among judges (eg, people, instruments) when multiple participants are measured and a true value is not known.¹⁵⁻¹⁷ Intraclass correlation coefficients reflect the ratio of judge variance to participant variance.¹⁵⁻¹⁸ If participant variance is low, the ratio is skewed, and a negative or low ICC value results despite low judge variance.¹⁶ Our participant variance was 0 because we were measuring stable water baths. Calculations of ICCs for our data would result in extremely low or negative ICC values despite temperature differences of less than 0.05°C.¹⁶

Another issue with using ICCs to determine the reliability of our temperature data is that ICCs report the amount of agreement or consistency among judges.^{16,17} Our research question was “How different are temperature measurements from a known criterion standard (ie, a NIST-calibrated mercury thermometer certified to 0.1°C)?” rather than “How consistent are thermocouples with one another?” If we assessed the consistency of the mercury thermometer and thermocouples, which would dismiss the idea that we know how accurate the mercury thermometer is, we would still have low ICC values because of the low participant variability.

CONCLUSIONS

Investigators examining temperature changes with therapeutic modalities should be cautious when using IT-21 thermocouples regardless of the electrothermometer used and when using IT-18 thermocouples for temperatures less than 10°C. It may also be helpful for investigators to record the amount of use each thermocouple receives, especially for the IT-21s. We recommend that investigators measure and report the validity, reliability, and calculated uncertainty (validity + reliability) of their temperature measurements. These validity and reliability measurements should be taken at temperatures similar to those expected during the experiment. This information will enable investigators and clinicians to better interpret results and conclusions.

REFERENCES

1. Jutte LS, Knight KL, Long BC. Reliability and validity of electrothermometers and associated thermocouples. *J Sport Rehabil.* 2008;17(1):50-59.
2. Jutte LS, Knight KL, Long BC, Hawkins JR, Schulthies SS, Dalley EB. The uncertainty (reliability and validity) of three electrothermometers in therapeutic modality research. *J Athl Train.* 2005; 40(3):207-210.
3. *Stedman's Medical Dictionary for the Health Professions and Nursing.* 5th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2005:1539.
4. Mish FC. *Merriam-Webster's Collegiate Dictionary.* 11th ed. Springfield, MO: Merriam-Webster; 2003:1051.

5. Ramsey FL, Schafer DW. *The Statistical Sleuth: A Course in Methods of Data Analysis*. 2nd ed. Pacific Grove, CA: Duxbury Press; 2002:102.
6. Otte JW, Merrick MA, Ingersoll CD, Cordova ML. Subcutaneous adipose tissue thickness alters cooling time during cryotherapy. *Arch Phys Med Rehabil*. 2002;83(11):1501–1505.
7. Myrer WJ, Myrer KA, Measom GJ, Fellingham GW, Evers SL. Muscle temperature is affected by overlying adipose when cryotherapy is administered. *J Athl Train*. 2001;36(1):32–36.
8. Myrer JW, Measom GJ, Fellingham GW. Temperature changes in the human leg during and after two methods of cryotherapy. *J Athl Train*. 1998;33(1):25–29.
9. Long BC, Cordova ML, Brucker JB, Demchak TJ, Stone MB. Exercise and quadriceps muscle cooling time. *J Athl Train*. 2005;40(4):260–263.
10. Columbus Instruments International Corporation. *Iso-Thermex Electronic Thermometer Instruction Manual*. Columbus, OH: Columbus Instruments International Corporation; 1992.
11. Commtest Instruments Ltd. Appendix specifications. In: *MMS3000-T6V4 Owner's Manual*. Christchurch, New Zealand: Commtest Instruments Ltd; 2001:75.
12. Physitemp Instruments. Temperature microprobes. In: *Physitemp Precision Temperature Specialists*. Clifton, NJ: Physitemp Instruments, Inc; 1999:4.
13. Merrick MA, Jutte LS, Smith ME. Cold modalities with different thermodynamic properties produce different surface and intramuscular temperatures. *J Athl Train*. 2003;38(1):28–33.
14. Merrick MA, Bernard KD, Devor ST, Williams MJ. Identical 3-MHz ultrasound treatments with different devices produce different intramuscular temperatures. *J Orthop Sports Phys Ther*. 2003;33(7):379–385.
15. Denegar CR, Ball DW. Assessing reliability and precision of measurement: an introduction to intraclass correlations and standard error of measurement. *J Sport Rehabil*. 1993;2(1):35–42.
16. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil*. 1998;12(3):187–199.
17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420–428.
18. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *The Statistician*. 1983;32(3):307–317.

Address correspondence to Blaine C. Long, PhD, ATC, LAT, Department of Health and Human Performance, 180 Colvin Recreation Center, Oklahoma State University, Stillwater, OK. Address e-mail to blaine.long@okstate.edu.