

SERIE: ¿CÓMO SE COMPORTA LA INVESTIGACIÓN EDUCATIVA EN CUBA, CINCO AÑOS DESPUÉS?

NÚMERO No.8: NO SIEMPRE LA 'MUESTRA' ES LA MUESTRA, NI SE PUEDEN REALIZAR GENERALIZACIONES A PARTIR DE ELLA

Dr. Cs. Paul A. Torres Fernández
Profesor e Investigador Titular, ICCP

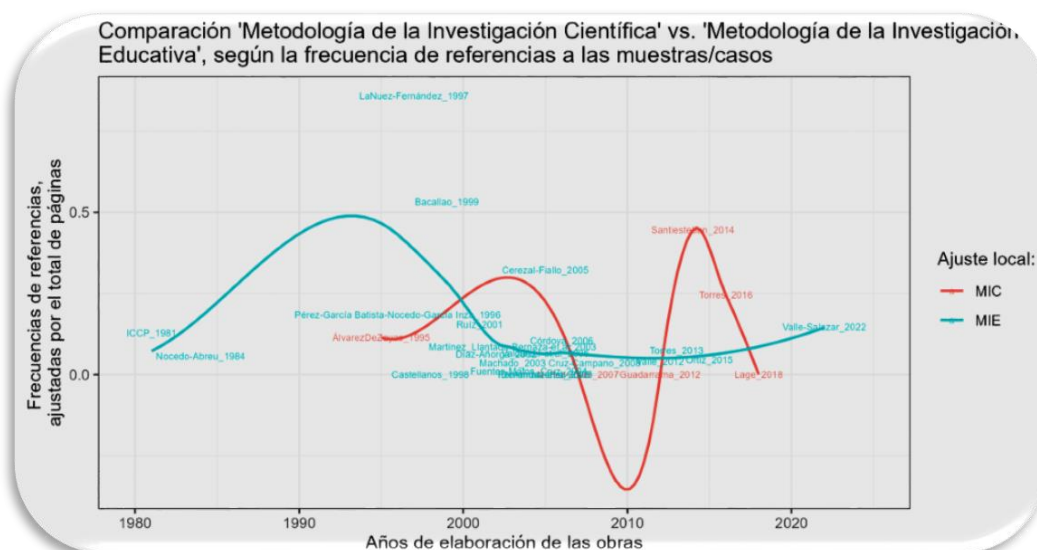
orcid.org/0000-0002-7862-2737

[linkedin.com/in/paul-antonio-torres-fernandez-4684048a](https://www.linkedin.com/in/paul-antonio-torres-fernandez-4684048a)

<https://www.facebook.com/ptorresfernandez>

<https://twitter.com/paintelectual>

<https://paulantoniotorresfernandez.blogspot.com>



12.06.2022; 17:51hs.

Introducción

“No toda persona puede dominar los fundamentos teóricos y los complejos instrumentos de la ciencia de hoy, pero debiéramos lograr que la cultura dotase a todas las personas instruidas de una capacidad para distinguir la ciencia de las pseudociencias y no dejarse confundir (...) El cubano despierto y culto, que aspiramos surja de nuestro sistema educacional, debe ser capaz de ver la diferencia ante cada propuesta concreta. La refutabilidad, la capacidad predictiva y la verificación son las marcas distintivas de una idea científica (...) A otras cosas, búsquese otro nombre, pero ciencia no son” (Lage, 2018; pp.121-122).

El tema de hoy –la delimitación de la ‘muestra’– es uno de esos aspectos de esta Serie que mucho tiene que ver con la alerta (¡y justificada exhortación!) que nos hace ese excelso científico cubano contemporáneo y líder indiscutible de *proyectos de investigación*; más aún, conductor de instituciones científicas cubanas con la probada capacidad de generar grandes y tangibles *impactos* sociales: el Dr. A. Lage Dávila.

Seré sincero; de los 1377 resultados de investigación educativa revisados con vistas a (Torres, 2016), fueron justamente el tratamiento de la ‘muestra’ y de la Estadística

aplicada (foco de nuestros siguientes posts) las áreas donde más falencias y errores conceptuales aberrantes pude apreciar.

Por eso, sin pretender ofrecer conferencias ni abarcar la imprescindible profundidad de la literatura especializada en estos temas, haré mi máximo esfuerzo por transmitir nociones claras y sugerir procedimientos de trabajo de la manera más elemental y accesible posible. Por suerte, no estamos en los años '80 ni los '90 y la Informática ha avanzado muchísimo (incluso para nuestro bloqueado país), y tenemos acceso a *software libres* y *amigables* que puedo sugerirles.

Claro; antes repasemos una vez más nuestra ya tradicional Figura No.2 para recordar que estamos tratando de reflexionar en torno a puntos clave de un proceso netamente *subjetivo* (configurado y desarrollado esencialmente en nuestras mentes), que tiene por delante el desafío enorme de '*penetrar*' el *objeto de estudio* y aprehender su *esencia*, o sea las *fuerzas internas* de su desarrollo *objetivo*. La *epistemología* de la *dialéctica materialista* es nuestra gran aliada en ese empeño y todas sus propuestas en esa dirección pueden encerrarse en una idea básica: ¡buscar la mayor *objetividad* posible, en cada paso del proceso investigativo!

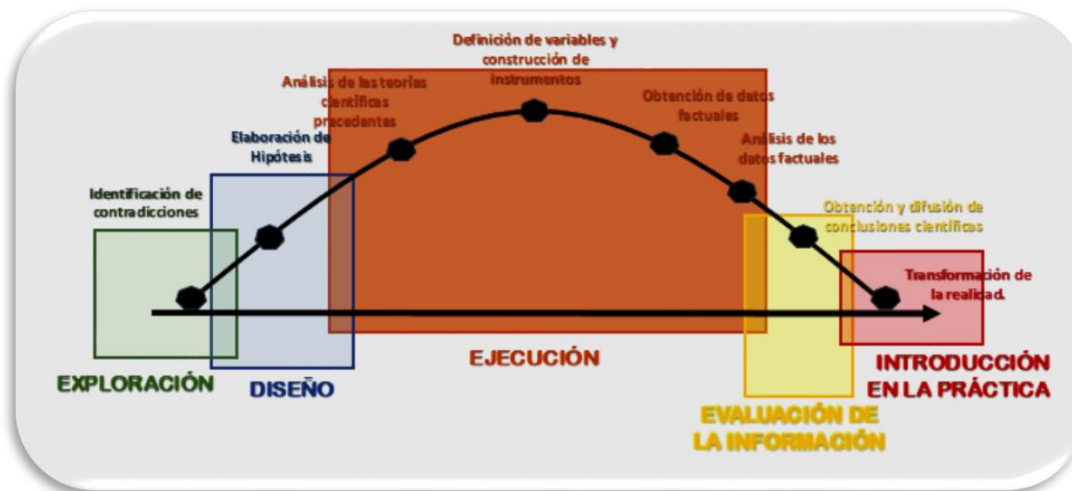


Figura No.2: Representación del carácter *subjetivo* del acto investigativo, en contraste con la naturaleza *objetiva* de la práctica. (Elaboración propia)

Para adherirnos firmemente a ese precepto –ya lo hemos visto– nuestro proceder científico debe transitar por las siguientes acciones secuenciales:

1. explorar en la práctica *contradicciones dialécticas* asociadas al *objeto de estudio* (concretamente, *desequilibrios* entre la manifestación de su *estado ideal*, por un lado, y del *real*, por el otro);
2. formular un *problema científico* (que precise una de esas *contradicciones*), y del cual se deriven: el *objeto de estudio*, el *objetivo de investigación* e *hipótesis científicas* (estas últimas, repuestas anticipadas y esperadas de aquel);
3. conformar un *marco teórico-referencial* que delimite los referentes teóricos conocidos, entre los previamente desarrollados por la ciencia en torno al *objeto de estudio*, de manera que se sometan a crítica científica los más *débiles* y, de forma fundamentada, se resalten los más *convincentes*, incluyendo –eventualmente– la elaboración propia de otros;

4. precisar las *variables principales* (alternativamente: *categorías*, en el caso de la modalidad de '*teoría fundamentada*' del *enfoque cualitativo*), presentes en los juicios que se establecen en las *hipótesis de investigación*, y de alguna manera presentes también en el *objetivo de investigación* y en la formulación del *problema científico*;
5. establecer las *definiciones conceptual y operacional* de esas *variables principales*, con lo cual se establece –a la vez– el '*camino descendente*' hacia los *datos* (o *testimonios/observaciones*, por ejemplo, en la modalidad '*etnográfica*' del *enfoque cualitativo*), a la vez que se crean las condiciones para el empleo futuro del '*camino ascendente*';
6. construir los *reactivos* de los *instrumentos de investigación*, de manera que se asocie a cada *indicador* un único *reactivo* y, una vez dispuestos todos en aquellos, se verifique la *validez y confiabilidad* de esas herramientas de captura de *datos* (o *testimonios/observaciones*) en la práctica misma;
7. desarrollar el *trabajo de campo* a través de una aplicación cuidadosa y efectiva de los *instrumentos de investigación*, a la vez que una recogida rigurosamente ordenada de los *datos* (o *testimonios/observaciones*) que dicha aplicación genere;
8. analizar esos *insumos factuales* con apego al aprovechamiento óptimo del '*camino ascendente*' (es decir, primero valorar individualmente cada uno de los *indicadores* de una misma *dimensión*, para después juzgar el estado de la dimensión que los contiene y, por último, enjuiciar el comportamiento de las *variables principales* a partir de lo concluido en torno a cada una de sus *dimensiones* respectivas);
9. emitir un juicio valorativo final sobre la *validez* (o *rechazo* parcial o total) de las *hipótesis de investigación* prefijadas al inicio o durante el proceso investigativo; y
10. elaborar *conclusiones y recomendaciones*, como resultado de lo anterior.

Claro, nos hemos estado refiriendo aquí solo a las acciones asociadas a la relación *objeto de estudio–investigador(es)*, pues también se deben ejecutar otras desde la perspectiva *sociológica* (o sea, centradas en la relación: *investigadores–pares–sociedad*) para poder hablar –en rigor– de un proceso investigativo completo.

Es importante recordar además que sugerí (en los marcos de esta Serie de posts) asumir la construcción (quizás, mejor *modelación*) de la *solución teórica* del *problema científico* (para el caso de las '*investigaciones explicativas*', las más frecuentes en la práctica comunitaria) como una extensión de la conformación del *marco teórico-referencial* (paso 3), dada su singularidad dentro del amplio espectro de *tipos y modalidades* de investigaciones posibles.

Asimismo debo reconocer, en relación con el procedimiento secuencial arriba descrito, que soy consciente de que hacia el interior de nuestra comunidad científica nacional se ha extendido la práctica (desde mi punto de vista, epistemológicamente errada) de intercalar entre el *diseño de la investigación* y la conformación del *marco teórico-referencial* un '*diagnóstico*' del comportamiento del *objeto de estudio* en la '*muestra*' asumida.

Pero prefiero posponer el análisis crítico de este proceder para un próximo post, en el contexto de mis reflexiones sobre el empleo de la Estadística aplicada en la investigación educativa cubana, que es el tema en torno al cual –intuyo– se generó esta atipicidad muy nuestra.

Aclarando conceptos básicos: la ‘muestra’ no es siempre la muestra

Entrando ya en el objeto de análisis del presente post, debemos recordar –ante todo– que el *concepto de muestra* es una categoría indiscutible de la Estadística, si bien es cierto que –a fuerza del predominio del criterio de un único ‘*método científico*’, que originó el prolongado ‘reinado’ del *enfoque cuantitativo*– se ha estado manejando esta como un componente más del proceder metodológico; es decir, parece haber sido absorbido este por la Metodología de la Investigación Científica.

Pero ya vimos (en nuestro tercer post) que no debiera hablarse –en rigor– ‘*del método científico*’, sino ‘*de los métodos científicos*’ (mejor, *enfoques investigativos*). De modo que, pensando en estos últimos, encontramos una razón más para rechazar la *usurpación* de la categoría ‘*muestra*’ a la Estadística, por parte de la Metodología de la Investigación Científica.

Desde mi modesto punto de vista, esa denominación debiera sustituirse –en esta otra disciplina científica– por la de ‘*unidades de análisis*’, como *definiendum* (término con que se define) de un *concepto superior* que incluye a los de ‘*muestra estadística*’ (para las modalidades del *enfoque cuantitativo* que lo demanden) y de ‘*casos*’ (para las modalidades del *enfoque cualitativo*), tal y como asumí en (Torres, 2016).

No es un preciosismo voluntarioso, tiene un fundamento lógico. En las modalidades del *enfoque cualitativo*, donde no se pretende hacer inferencias estadísticamente rigurosas a conglomerados más amplios que los estudiados, no cabe el concepto de *muestra*, pues ella supone el aseguramiento previo de condicionantes, como los criterios de *aleatoriedad* y de *representatividad* de la *población* (su par dialéctico).

En su lugar, la denominación de ‘*casos*’ la libera de esas condiciones, para dejar en su lugar las que representen al *objeto de estudio*, que es la única condición (*definiens*) –a su vez– del *concepto superior* de ‘*unidades de análisis*’ propuesto. Dicho más claramente, en los marcos de la Metodología de la Investigación Científica debiéramos hablar –mejor– de ‘*unidades de análisis*’ en general, pudiendo estas constituir ‘*casos*’ (para el *enfoque cualitativo*) o ‘*muestras estadísticas*’ (en los marcos del *enfoque cuantitativo*). Eso nos evitaría varios errores metodológicos.

Por ejemplo, en la Obra No.505 –entre las consultadas para (Torres, 2016)– se afirma: “*Para seleccionar la muestra se utilizó la técnica probabilística con un criterio intencional que responde a un interés diatópico, teniendo en cuenta la procedencia de los estudiantes de la zona urbana y la zona rural*” (p.18). Sin embargo, por un lado, ‘*intencional*’ no es un sinónimo, sino un antónimo de ‘*probabilístico*’; por el otro, la *muestra estadística* debe ser necesariamente *aleatoria*. El concepto de ‘*muestra intencional*’ no existe para constituirse en una opción procedimental de la Estadística Inferencial, sino para alertar de *algo* que no se debe utilizar en ella.

No es un caso aislado; una antonimia similar se presenta en la Obra No.1222, cuando se plantea: “*El muestreo es probabilístico intencional, estratificado y aleatorio, que representa a la población en sexo, edad, municipios*” (p.15). Esas *unidades de análisis* son, por tanto, *casos* y no *unidades muestrales*.

Otro tipo de error asociado a los conceptos de ‘*unidades de análisis*’ y de ‘*muestra*’ se apreció en la Obra No.77, cuando se afirma: “*Como población se consideran los dirigentes de las 33 escuelas politécnicas de Ciudad de La Habana. La muestra está representada por todos sus municipios y las especialidades que se estudian en ellos [sic]*” (p. 87).

Más que un problema de redacción, se ha obviado que las '*unidades muestrales*' (en tanto '*unidades de análisis*') deben constituir representantes del *objeto de estudio*. En el reporte de investigación el *objeto de estudio* declarado fue '*la capacitación [sic]*' para el "*desarrollo de la competencia comunicativa en los dirigentes de la Educación Técnica Profesional (ETP)*".

Los portadores de esa '*competencia*' (que se pretendía fortalecer con el estudio científico practicado) eran los dirigentes educacionales de la ETP de la capital del país, no las estructuras organizacionales en que las que ellos se ubicaban (es decir, las especialidades de la ETP y las Direcciones Municipales de Educación). No debe perderse de vista la noción primaria de representación del *objeto de estudio* (lo mismo en la *muestra* que en los *casos*), pues es en el *objeto* donde se define el atributo de interés que deben portar aquellos.

Aprovecho este análisis para cuestionar otro pseudo concepto que he apreciado en el desarrollo de la '*cultura de los tribunales*'. En ese sentido, es importante resaltar que no necesariamente las *unidades de análisis* tienen que ser sujetos. Al parecer, el manejo del término '*población*' a la par del de '*muestra*' ha sembrado en algunos investigadores la idea de que ambos se refieren siempre a individuos.

El atributo delimitado por el *objeto de estudio* de una investigación puede ser portado en representación de él, por ejemplo, por un grupo de *planes de estudio* (en los marcos de estudios curriculares), o por *hechos históricos* (para investigaciones de ese carácter), o por *medios de enseñanza* (en indagaciones pedagógicas dirigidas especialmente a esa categoría didáctica), entre otras posibilidades. De ahí la pertinencia del *concepto superior* de '*unidades de análisis*' que he defendido aquí.

Por último, quiero añadir a estas reflexiones sobre los conceptos relacionados con el de *muestra* otro tipo de sesgo identificado también entre los 1377 reportes de investigación revisados para (Torres, 2016). Se trata de hablar de *muestra* aun cuando ella termina coincidiendo con la *población*. Por ejemplo, en la Obra No. 911 se plantea:

"Muestra: 30 profesores que imparten la Educación Laboral en el 9no. grado en la ESBU"(...). La población y la muestra en este caso coinciden, teniendo en cuenta que con el nuevo modelo educativo, los Profesores Generales Integrales deben rotar por todas las asignaturas" (p.13).

Asimismo, aunque correspondiente a otro territorio del país, en la Obra No.931 se señala de forma análoga: "*Se tomó como población 16 bibliotecarios de la Educación de Adultos del municipio (...). La muestra coincide con la población (...)*" (p. 14).

Si bien, en principio, la idea de igualar la *muestra* a la *población* es teóricamente aceptable, no debiéramos perder de vista que sugerir el manejo futuro de una *muestra*, en los marcos de un trabajo de investigación científica, significa reconocer que se pretende realizar –en algún momento– una *inferencia estadística* de las medidas obtenidas en la *muestra* a toda la *población* (con cierto margen de error, pues no existe ninguna *inferencia estadística* absolutamente segura). Es suficiente hacer referencia a que se realizará un *estudio censal* y no crear la expectativa.

Dejaré pendiente para un próximo post la explicación, en detalles, del término aquí empleado de '*inferencia estadística*', una vez aseguradas otras nociones básicas, más elementales.

Más que el tamaño de la muestra, lo esencial radica en otras condicionantes

Otra tendencia también identificada en torno al trabajo con las *muestras*, desde la experiencia profesional (oponente y miembro de tribunales de tesis, revisor de proyectos de investigación, jefe de programas de investigación, etc.) es la de prestarle mucha atención al cálculo de su tamaño.

Y no es que eso esté mal; es legítimo preocuparse por considerar solo una cantidad razonable de *unidades de análisis*, de manera que no se eleven los costos del estudio innecesariamente, o se prolongue el *trabajo de campo* por mucho tiempo, etc. La dificultad viene cuando se piensa solo en esa condición, obviando otras más trascendentes en el orden estadístico y metodológico.

En la literatura especializada existe una abundante información sobre cómo calcular el *tamaño* –digamos, *óptimo*– de una *muestra* (por ejemplo, parece comprensible hasta para los noveles la propuesta del sitio-Web ‘SurveyMonke’ y su calculadora, disponibles en: <https://es.surveymonkey.com/mp/sample-size-calculator/>). También pueden encontrarse buenas recomendaciones de *tamaños de muestra*, según el tipo de estudio, e incluso ejemplos desarrollados sobre su cálculo, en la muy referida obra (Hernández-Sampieri, Fernández & Baptista, 2010).

Básicamente, se requiere combinar en las fórmulas estadísticas cinco elementos: (i) el tamaño de la *población* de la cual se va a extraer la *muestra*, (ii) la *probabilidad* de que el *evento suceda*, (iii) el *margen de error* que se está dispuesto a asumir, (iv) el *nivel de confianza* en los resultados que se obtendrían a partir de ella y (v) el puntaje correspondiente a ese *nivel de confianza* en la *distribución teórica de probabilidades* asociada. Pero, ¡tranquilos!... No pienso presentar ninguna fórmula matemática ni estadística en mis posts, pues prefiero retener a mis lectores.

Una derivación lógica de la explicación anterior (y aún sin tener a la vista la fórmula) es que el *tamaño óptimo* de una *muestra* NO es proporcional al *tamaño de la población* de donde se va a extraer, pues no solo depende del *tamaño de la población*. Más claramente, ¡el *tamaño óptimo de una muestra* no es un por ciento del *tamaño de su población*!

Este es otro error conceptual frecuentemente cometido por nuestros investigadores (y sus conductores), de acuerdo con lo estudiado con vistas a (Torres, 2016). Un claro ejemplo de ello se presenta en la Obra No. 1012; el autor (con la autorización de sus tutores) señala: “*Para desarrollar la evaluación en los diferentes centros las muestras serán seleccionadas de una manera simple aleatoria y en ningún caso estarán constituidas por menos del 10% de la población*” (p. 58).

Por su parte, en la Obra No. 1207 se señala en relación con este tema:

“*La muestra que se utilizó está aproximadamente en correspondencia con el 10% de los obreros y el 12,6% de los técnicos en formación de ambos Centros Politécnicos, el 10% de los directivos de la esfera de la producción; el 100% de los profesores que trabajan con la especialidad en el Pelotón Docente y el 100% de los dirigentes y metodólogo de la especialidad*” (p. 58).

Ahora bien, más importante que el *tamaño de la muestra* es asegurarse de que la que se extraerá sea totalmente *aleatoria* (¡y nunca *intencional*!); así como, que represente adecuadamente a su *población*. La cuestión de la *aleatoriedad* nos resulta más familiar a todos, por el aquello de que son populares ciertos recursos elementales de

extracción (como la *lotería*, los *números aleatorios* generados por computadoras, etc.). Sin embargo, es mucho menos conocida la cuestión de la *representatividad*.

Trataré de reproducir un ejemplo ilustrativo que, casualmente, me vi en la necesidad de improvisar en un análisis no previsto sobre el tema, recientemente. Supongamos que queremos estudiar el nivel de satisfacción de los pasajeros de ómnibus urbanos de una determinada ciudad. Después de verificar la *validez* de la *guía de entrevista* que hemos elaborado, pasamos a realizar un *estudio piloto* con ella, para revisar el nivel de *confiabilidad* del instrumento, como vimos en el post anterior.

Para efectuar ese estudio previo necesitamos extraer una *muestra* de la *población objetivo* (aquí, todos los potenciales pasajeros de los viajes de ómnibus de la ciudad). En principio, y de acuerdo con la tendencia arriba referida, decidimos entrevistar a los pasajeros de cierto ómnibus urbano, durante uno de sus recorridos diurnos, de manera que se encuentren situados de la mitad hacia delante del ómnibus. ¡Se trata de una *muestra grande*, el 50% de los pasajeros del ómnibus al momento de la entrevista, e incluso elegimos ese ómnibus al azar!

Sin embargo, muy probablemente los resultados de ese *estudio piloto* nos confundirá con relación a la *aplicación definitiva* de la *guía de entrevista*, puesto que no se han tenido en cuenta varias condiciones cambiantes del *piloto* a la *aplicación definitiva*: hacia la parte delantera se encuentran los asientos de las embarazadas e impedidos físicos, esta parte del ómnibus dispone de más esparcimiento para los pasajeros que viajen de pie (por lo que harán un viaje relativamente más cómodo), no es el mismo tipo de pasajero el que viaje por el día que el que lo hace por la noche y la madrugada, no en todos los puntos del recorrido es la misma afluencia (pudiendo en algunos lugares generarse mayores niveles insatisfacción que en otros), etc. Es decir, la verificación del criterio de *representatividad* de la *población* es tan importante como la selección *aleatoria* de la *muestra* (¡y más que su *tamaño*!).

Claro, un conocedor de las nociones básicas de la '*teoría del muestreo*' pudiera interpellarme y señalame que, de acuerdo con el '*Teorema del límite central*', la *selección aleatoria* de 100 o más *unidades de análisis* "(...) *será una muestra con una distribución normal en sus características*, [por] *lo cual sirve para (...) hacer estadística inferencial*" (Hernández-Sampieri, Fernández & Baptista, 2010, p. 189).

Y sí, es cierto. Pero una '*distribución normal*' es una *distribución acampanada* y, por tanto, con dos *colas*. Luego, aun eligiendo de forma totalmente *aleatoria* una *muestra*, existen *probabilidades* (pocas, pero existen) de que presente características *límites* ('*anormales*'), y resulten *unidades muestrales* con un comportamiento *muy bueno* en el *atributo* de interés; o, por el contrario, con comportamiento *muy malo* (Figura No.3).



Figura No.3: Ejemplo de distribución normal (Tomado de: Hernández-Sampieri, Fernández & Baptista, 2010, p. 308).

No estoy hablando de cuestiones casi improbables. En ocasión de la realización de uno de los Operativos Nacionales de Evaluación de la Calidad de la Educación, realizado en el país a mediados de la década del 2000, la *muestra* de escuelas de la capital se recargó –con apego a la *aleatoriedad*– en sus tres mejores municipios de entonces, entre los cinco que representaron a la provincia en ese evento evaluativo.

Como era de esperar, los resultados en las pruebas aplicadas se elevaron con relación al comportamiento histórico, creándose altas expectativas entre los dirigentes educacionales de la capital y del Ministerio. Sin embargo, al año siguiente la *muestra* resultante para la provincia volvió a asumir una tendencia a la *normalidad* (en términos de *representatividad*) y los resultados de las pruebas de aprendizaje descendieron. ¡Así de dramático puede ser el tema del *muestreo* en los estudios de investigación científica, si no se fortalece el *criterio de la representatividad* de la *población*!

Entonces, ¿qué hacer para obtener una ‘buena muestra’?

Bueno, lo primero ya lo dijimos: si la investigación se guía por el *enfoque cualitativo* o, aun siguiendo el *enfoque cuantitativo*, el análisis de sus resultados se limitaran a una *población* razonablemente pequeña, entonces lo mejor es no hablar de *muestra*, sino de *casos* o de *estudio censal*, correspondientemente.

En otras situaciones (probablemente para *estudios explicativos, correlacionales o explicativos* que se orientan por el *enfoque cuantitativo* de investigación, y que tienen pretensiones de realizar las llamadas ‘*generalizaciones*’), se debe elegir una *muestra* suficientemente grande (o sea, de alrededor 100 o más *unidades de análisis* o, en caso extremo, de no menos de 30) y verificar durante el *estudio piloto* que se han tenido suficientemente en cuenta aquellas características de la *población objetivo* que pudieran debilitar el cumplimiento del *criterio de confiabilidad* de los *instrumentos*.

Pero si el proyecto de investigación que tendrá a cargo el estudio dispone de tiempo, y de la capacidad técnica y financiera requerida, o el nivel de responsabilidad de sus futuros resultados científicos en la toma de decisiones es muy elevada, entonces lo correcto sería realizar un ‘*muestreo riguroso*’. La pregunta que seguro sigue es: “¿y qué es un ‘*muestreo riguroso*’?”.

Pues, acudir a un ‘*muestreo complejo*’, que es el que se utiliza –por ejemplo– en los llamados ‘*estudios a gran escala*’ (como PISA, TIMSS, PIRLS, ERCE, etc.). Ya la propia denominación nos da una idea del grado de rigor técnico de este otro proceder. El *muestreo complejo* parte de la premisa de que para que un *muestreo aleatorio simple* asegure, además de la *aleatoriedad*, una *representación* adecuada de la *población* se requiere de no menos de 50 mil *unidades muestrales*.

Ello significaría, como es de esperar, costos muy altos para el estudio. Por ejemplo, en el caso del recientemente realizado ERCE-2019 (Torres, 2020a) se ocuparon en Cuba tan solo 250 escuelas primarias y poco más de 5 mil estudiantes por cada uno de los dos grados evaluados. Mas, se estaría requiriendo para un *muestreo aleatorio simple* con garantía de *representatividad* 10 veces más estudiantes por grado y más de 2 mil instituciones escolares, en lugar de aquellas cantidades realmente utilizadas.

De modo que en el ‘*muestreo complejo*’ la prioridad es fortalecer la *representatividad*, aunque ello produzca un ‘*daño*’ a la *aleatoriedad* (dado que solo el *muestreo aleatorio simple* asegura que todas las *unidades de análisis* tengan la misma *probabilidad* de ser elegidas).

Ese debilitamiento consciente de la *aleatoriedad* se lleva a efecto a través de un proceso de *muestreo aleatorio* pero *por etapas*. En el caso del ERCE-2019, como mismo antes en el TERCE-2013 (OREALC-UNESCO, 2016), se utilizó un *muestreo bi-etápico*; pues aun cuando eran los estudiantes los referentes básicos del *objeto de estudio* y, por tanto las *unidades de análisis*, primero se le prestó atención a la *representatividad* de los distintos tipos de escuelas donde estaban ellos matriculados.

Las escuelas son ubicadas según los *estratos* (es decir, distintas agrupaciones de las *unidades de análisis* de acuerdo con diferentes opciones de los *criterios* poblacionales preestablecidos). En esos dos estudios evaluativos regionales se asumieron 12 *estratos*, que resultaron de las distintas combinaciones generadas al relacionar tres *criterios*: (i) *dependencia* (pública o privada), (ii) *área* (urbana o rural) y (iii) *grados en la escuela* (con solo tercer grado, con solo sexto grado, o con tercero y sexto grados); claro, para Cuba *colapsaron* los *estratos* donde intervenía la *dependencia privada*.

El *muestreo estratificado* es obviamente también *aleatorio* (hacia el interior de cada *estrato*), pero dispone de tres formas de '*afijación*' de las *muestras* (Torres, 2020b): (i) *uniforme* [es decir, asignando igual cantidad de *unidades de análisis* para cada *estrato*], (ii) *proporcional a los tamaños de los estratos*, y por último (iii) *estratificado óptimo* [o sea, atendiendo a la *desviación estándar* de la *variable* de interés en cada *estrato*]. En los dos estudios evaluativos regionales arriba referidos se optó por un *muestreo proporcional al tamaño* (de los *estratos*).

Una vez seleccionadas las escuelas (y vencida la primera *etapa* del *muestreo*), se aprovechó su estructura organizacional y se seleccionó *aleatoriamente* un aula (o salón de clases) por grado en cada escuela. En esta otra *etapa* se realizó, entonces, un *muestreo por conglomerados*, de manera que participaron –en principio– todos los estudiantes de cada aula (*conglomerado*) elegida. Así se completó la *muestra*.

Ahora bien, ya explicamos que la decisión de partir de priorizar la *representatividad* creó un '*daño*' en la *aleatoriedad*, '*daño*' que hay que corregir antes del procesamiento estadístico final de los *datos* que posteriormente se obtendrán, durante el *trabajo de campo*. Es por ello que, una vez efectuado este último y precisada la *muestra efectiva* del estudio se pasa a calcular y utilizar los llamados *pesos muestrales*.

Los *pesos muestrales* se calculan como el *recíproco* de la *probabilidad* de haber sido elegido un algún paso del proceso. Así, por ejemplo, si una escuela disponía de tres aulas, entonces la *probabilidad* (p) de que fuera elegida su aula es $1/3$ y el peso (w) de sus estudiantes es 3; para una escuela con 5 aulas, el peso de cada alumno es 5.

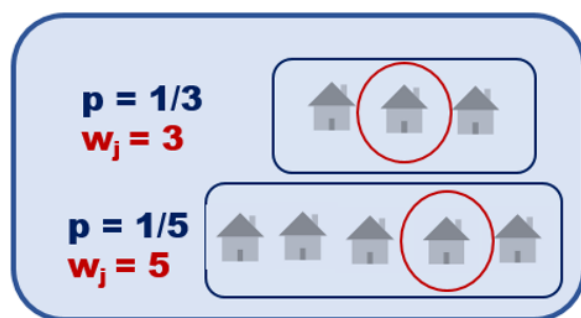


Figura No.4: Ejemplo de cálculo del peso del estudiante (Elaboración propia).

Y es que no es lo mismo *representar* en la *muestra* un aula entre tres candidatas de ser elegidas, que entre cinco. De modo que las puntuaciones estimadas de cada uno de los estudiantes de la primera escuela deben multiplicarse por tres, mientras que en la segunda debe hacerse por cinco.

En rigor, este proceso de determinación del valor por el cual hay que '*balancear*' las puntuaciones estimadas de cada estudiante participante en un estudio, en el que se utilizó un '*muestreo complejo*', involucra no uno sino varios *pesos muestrales*: el *peso de escuela*, el *peso de estudiante*, el *peso total* (que resulta de la *probabilidad combinada* de haber sido elegida su escuela y, a la vez, su aula), así como el llamado '*peso senatorial*', que intenta equilibrar las marcadas diferencias de poblaciones escolares entre los países participantes; por ejemplo entre República Dominicana y Brasil (OREALC-UNESCO, 2016).

El procedimiento de trabajo con las *muestras* en los estudios de investigación científica no concluye aquí, pues –en todo caso– estamos hablando de una '*muestra de ocasión*' (recuérdese las marcadas diferencias que se pueden presentar entre distintas *muestras* como resultado de su extracción, aun cuando sean todas de un mismo tamaño; como se ilustró en la Figura No.3).

Los potentes paquetes estadísticos con que se cuenta hoy (amparados en el enorme desarrollo alcanzado por las TIC) nos ofrecen la posibilidad de no depender solo de esa '*muestra de ocasión*' y de generar, sobre su base, otras muchas *muestras virtuales* que nos ayuden a *estimar*, con mayor precisión, las *medidas estadísticas poblacionales*, que resultarán imprescindibles para hablar de una '*generalización*' de los resultados del estudio, en sus conclusiones.

Existen diferentes técnicas para realizar ese útil proceso (*Jackknife*, *Linealización de Taylor*, *Bootstrap*, *Replicación Balanceada Repetida*, etc.), pero aunque responden al tema de hoy, las dejaremos para siguientes posts, una vez avanzado el análisis de la utilización de la Estadística aplicada a la investigación educativa, que iniciaremos en el próximo encuentro.

¡Los esperamos el próximo fin de semana! ¡Sigán con nosotros!

(Tomado de los Blogs 'Investigación Educativa en Cuba', de Google, y 'Evaluación Educativa', de CubaEduca)

Referencias bibliográficas

1. Lage, A. (2018). *La Osadía de la Ciencia*. La Habana, Cuba: Editorial Academia.
2. Hernández-Sampieri, R., Fernández, C. & Baptista, M. P. (2010). *Metodología de la investigación* (Quinta edición). México D. F., México: McGraw-Hill.
3. OREALC-UNESCO (2016). *Reporte Técnico. Tercer Estudio Comparativo y Explicativo*. Santiago, Chile: UNESCO.
4. Torres, P. (2016). *Retos de la investigación educativa actual. Aportes a su tratamiento*. Universidad en Ciencias Pedagógicas "Enrique José Varona". Cuba: La Habana (Recuperado de: <https://drive.google.com/file/d/17w13EvbTYYMR266KHI2UsFw9VDyCXwal/view?usp=sharing>).

5. _____ (2020a). Más del entorno informático R en la Investigación Educativa cubana. ¿Se puede predecir la muestra de Cuba en el ERCE-2019? (Primera parte). *Ciencias Pedagógicas*, 13(1), 201-2011. (Recuperado de: <http://www.cienciaspedagogicas.rimed.cu/index.php/ICCP/article/view/232>)
6. _____ (2020b). Más del entorno informático R en la Investigación Educativa cubana. ¿Se puede predecir la muestra de Cuba en el ERCE-2019? (Segunda parte). *Ciencias Pedagógicas*, 13(2), 84-91. (Recuperado de: <http://www.cienciaspedagogicas.rimed.cu/index.php/ICCP/article/view/244>)