

David Sulmont: “Muchas de las cosas que se dicen sobre las cifras tienen un sustento bastante débil”

Ideele Revista N° [233](#)

Por:

[María José Palacios](#)



(Foto: PUCP)

David Sulmont, actual director del Instituto de Opinión Pública de la PUCP, fue el encargado del área de Sistemas de Información de la CVR. Desde ahí veía todo lo concerniente a las estadísticas. Además fue miembro del Comité editorial del Informe Final. En esta entrevista responde de manera amplia acerca de algunos temas polémicos en torno a la CVR.

¿Cuál fue tu función específica dentro de la CVR?

Al principio, cuando empezamos, mi función era dirigir el diseño de la metodología para recoger los testimonios, todos los protocolos de entrevistas, los instrumentos, el diseño de la base de datos, los procedimientos para los análisis de testimonios y el registro de la información. Después, mi función fue también analizar la información ya registrada en la base de datos, las tendencias estadísticas, todo lo que eran los reportes estadísticos de la Comisión y todo lo que fue la reconstrucción de los casos que salían reportados en los testimonios.

¿Qué es lo que te gustaría destacar de esta experiencia?

Para empezar, el reto de reconstruir lo que ocurrió en esos años. En ese entonces, la CVR tenía un reto importante: tratar de llegar y darle voz a la mayor cantidad de personas posibles que hayan sido víctimas del conflicto armado interno. Teníamos el reto de recoger toda esa información y analizarla en menos de dos años.

Una de las cosas que creo que logramos es que todas las personas que dieron su testimonio en la CVR estén registradas, y los nombres de las personas y sus familiares estén registrados. Eso fue la base de datos de la Comisión de la Verdad que tiene cerca de 70 mil nombres de víctimas, de diferentes tipos de violación. Registrar todo eso en el tiempo y con los recursos que teníamos fue un reto importante.

Segundo, dar cuenta de lo que pasó a partir de lo que salía de los patrones de las violaciones reportadas y que han sido la base de las conclusiones de la Comisión de la Verdad respecto a la magnitud del conflicto que cuestionaban estimados previamente existentes y que hasta ahora siguen causando controversia. También el establecer, un poco, el nivel de responsabilidad que tenían los actores armados, sobre todo el caso Sendero Luminoso que resulta ser el perpetrador de la mayor cantidad de víctimas. Eso dice mucho de las características de la organización terrorista

¿Se imaginaron la reacción que iba a suscitar el informe?

Sí y no. El informe presenta una imagen muy dura del país y sabíamos que las personas que podrían verse interpeladas por eso eran, primero, los responsables de los crímenes y violaciones; y en segundo lugar, también a nivel político, las personas que justificaban cierta manera de actuar, tanto de parte del Estado como de Sendero luminoso. En algunos casos había adversarios muy poderosos en la política del país. Por eso, no es de extrañar que el informe haya tenido, en ciertos sectores, una reacción fuerte.

Uno de los enfoques centrales de la Comisión de la Verdad, que es característica de muchas comisiones de la verdad, es que está centrado en quienes fueron víctimas de violaciones, no en los proyectos políticos de las personas, de los actores e instituciones que participaron del conflicto. En ese sentido, hay una posición ética vinculada con el tema de la perspectiva del respeto de los Derechos Humanos. En el caso de la población más afectada, en el Perú son los menos poderosos y los que tienen poco peso para el país. Esa es una de las cosas que explica por qué la magnitud y por qué la sorpresa frente a esos estimados. Se pensaba que era menos.

Hay quienes cuestionan mucho las cifras de la CVR, incluso hablan de manipulación.

Sobre el tema de las cifras se ha dicho un montón de cosas, y muchas de las cosas que dicen tienen un sustento bastante débil.

Antes de la Comisión de la Verdad había algunos registros compilados por algunas instituciones. Por ejemplo, DESCO tenía una compilación de reportes vinculados a la violencia política que registraba en una base de datos y que luego publicó en un libro. Cuando uno comenzaba a sumar todos esos reportes, más o menos sale la cifra de los 25 mil.

Había también una base de datos de la Defensoría del Pueblo en el que se sistematizaba todos aquellos casos de denuncia por desaparición forzada. Todos esos registros eran con nombres de personas que figuraban en las denuncias como desaparecidas. En total eran entre 6 mil y 7 mil nombres. Estos no se pueden sumar a la de DESCO porque en los 25 mil no hay nombres en todos los casos y varios están repetidos o pueden estar repetidos. Además, una comisión del Senado en ese entonces había compilado una serie de estadísticas usando información de la Policía Nacional. Más o menos salían 12 mil personas en el período de los 80.

Por lo tanto, las fuentes que había antes eran de ese tipo en los que la mayoría de personas no estaban identificadas. Salvo los casos que están en la base de datos de la Defensoría del Pueblo, no había una lista de víctimas. Entonces la Comisión de la Verdad, con los testimonios, obtiene una lista de las víctimas reportadas en la que salen más o menos registradas unas 24 mil personas.

Pero eso es una enumeración de registro de personas, no es una estimación. Porque si uno dijera que solamente existen las víctimas que uno puede registrar estaría suponiendo que uno es capaz de registrar todo. Hay varios lugares donde la Comisión no pudo llegar por varios motivos, por recursos, tiempo, etc. Pretender que lo que una organización puede registrar es lo único que existe es un supuesto bastante endeble.

Lo que la Comisión hizo fue compilar los 3 registros que habían: los de la Comisión con víctimas identificadas; los de la Defensoría del Pueblo; y los de la Coordinadora Nacional de DDHH, la cual, durante la Comisión de la Verdad, se dio el trabajo de hacer un listado de todas las personas que aparecían en los expedientes de las instituciones que forman parte de la Coordinadora. Nos basamos en un trabajo que se había hecho en la Comisión de Guatemala: comparar las listas y ver cuánta gente aparece como repetida

entre las listas. Y eso fue lo que hicimos en el caso peruano. Teníamos tres listados y a través de una metodología que diseñamos y una serie de herramientas, hicimos el cruce de información de esas tres listas con la asesoría de Patrick Ball.

Para poder hacer el cruce de información había que preparar los datos, porque muchos reportes tenían la identificación incompleta de la gente. Por ejemplo, no todos los registros tenían apellido materno. En el caso de la Comisión de la Verdad, de los 24 mil registros de personas reportadas como muertas o desaparecidas, 18 mil tenían el nombre y los dos apellidos. Había unos 6 mil que no tenían ese nivel de información. Para que el cruce sea más exhaustivo, se trabajaba solamente con el registro que cumplía un mínimo de condiciones. Eso excluyó una serie de registros que estaban en las tres bases de datos. Se hizo la compilación y se encontró que en las tres bases de datos había un total de 24 mil personas reportadas. Para hacer el cruce, primero se verificaba si coincidían los nombres, los lugares y las fechas. Había toda una metodología para decidir si dos registros eran o no la misma persona, y eso se hacía además en dos equipos por separado. Hay toda una técnica que sirve para que los errores que uno puede cometer al hacer el cruce, se controle y se minimice. Eso tomó dos o tres meses haciendo eso.

A partir de esa compilación, se usó una técnica que viene de la estadística: la de captura y recaptura, en este caso, estimación de múltiples sistemas. A partir de la información que se tiene, se trata de estimar aproximadamente qué es lo que no se conoce. Ese es un principio básico de una estimación estadística. El universo de víctimas que hay no se conoce, no se sabe cuántas son. Se tienen tres listados que se cruzan y que son una parte del universo. Los 24 mil que hay no son todo el universo; porque, como dije anteriormente, de la Comisión de la Verdad sacamos 6 mil víctimas que no podíamos cruzar. Entonces, las 24 mil, son sólo una muestra de la población.

A partir de esa muestra se trata de estimar el tamaño de la población. Normalmente se usa una muestra probabilística, pero para eso se necesita un listado de la población, cosa que no teníamos. Entonces, lo que se utilizó fue una técnica del campo de la biología o la zoología, pero que después se ha usado en censos de población, una técnica que ahora es bastante utilizada en muchos campos.

La estimación de múltiples sistemas implica tratar de ver, a partir del análisis del traslape, qué hay entre las fuentes que tenemos y cuánto es lo que no se ha recogido. El trabajo de cruce de listas es identificar hasta qué punto estas listas se traslapan. Lo importante es identificar esta tasa de traslape, porque

la teoría nos dice que cuanto mayor es el traslape entre las listas, menor es el universo que no ha sido capturado o registrado.

El universo es como un cuarto cerrado, oscuro; y las listas, como tres personas que entran al cuarto a ciegas y tratan de ver qué tan grande es ese cuarto, pero no lo ven. Entonces comienzan a chocarse entre sí. Si la gente se choca mucho significa que el cuarto no es muy grande, porque no hay mucho espacio para recorrer; en cambio, si el cuarto fuera muy grande, la gente puede estar caminando mucho tiempo sin encontrarse con nadie. Hay una relación entre la cantidad de veces que tú te chocas con otra persona que está caminando a ciegas en el cuarto y el tamaño del cuarto. Más o menos esa es la lógica que se aplica a este problema. Cuanto más se traslapan las listas, implica que el espacio donde éstas se mueven es más pequeño. Cuanto menos se traslapan es más grande.

Aún así, para que eso se aplique hay una serie de problemas porque no se han recolectado aleatoriamente a las víctimas en cada lista. Éstas pueden tener un sesgo y, de hecho, lo tienen. Por ejemplo, en el caso de la Defensoría del Pueblo y las ONG, la mayoría de víctimas reportadas son cometidas por agentes del Estado. La CVR es la que reporta más víctimas de Sendero. Cada lista tiene un sesgo, pero el que se pueda comparar una lista con otra te permite modelar el sesgo. Modelar es representarlo en una ecuación matemática que te controla hacia dónde te jala más esa lista. Y la ventaja de tener tres listas es que permite múltiples controles. Dos listas pueden tener un sesgo, pero la tercera te controla.

Rendón siempre ha dicho que no es posible que Sendero Luminoso haya matado a más gente que el Estado, porque parte del supuesto que el Estado tiene más armas, pero esa no es una razón.

Además de esto, se hace lo que se llama una estratificación; es decir, se hace este mismo ejercicio de analizar cuánto se traslapan estas listas, pero para distintas regiones del país, pues el conflicto armado interno no fue igual en todas partes. Al identificar unas 57 regiones afectadas, partimos el país en 57 pedacitos; y en cada pedacito, se hizo esa estimación, y se hizo aparte para Sendero, aparte para el Estado. La cantidad de ecuaciones que se construyen para hacer eso es como 59 a la séptima potencia, es un número enorme de ecuaciones. Después es necesaria una serie de pruebas estadísticas para hacer lo que se llama el nivel de ajuste de estos modelos y se escoge aquellos que, para regiones específicas, capturan mejor la dinámica entre los proyectos que

documentan el conflicto armado interno. De ahí salen estimados que van sumándose y se llega a la cifra total que es un aproximado. Al ser una estimación estadística, tiene un margen de error. En el caso de la Comisión va de 54 mil a 70 mil. La conclusión es que tenemos un 95% de confianza de que la cifra de víctimas está en algún punto entre estos dos límites, por eso se dice que está alrededor de 70 mil personas.

Hay un montón de teoría, estadísticas, matemáticas y computación detrás de esto. En muchos proyectos usan esta metodología para distintos problemas que dan cuenta de que los estimados que se producen usando este tipo de métodos son robustos. Es decir, efectivamente apuntan hacia las dimensiones del problema que se quiere estimar. Un ejemplo es la estimación de la gente con VIH. En algunos proyectos lo que hacen es analizar las listas de instituciones que registran este tipo de población. Entonces van comparando y van haciendo análisis de traslape y tienen un estimado. En el caso de un censo, se debe registrar a toda la población que hay en un país. Todo censo tiene una tasa de no cobertura. En el caso peruano a veces es difícil registrar a toda la gente que vive en la Amazonía o en zonas más alejadas.

Entonces, ¿cómo calculas la gente que el censo no ha registrado? Por definición, el censo registra a todo el mundo, pero en la práctica uno sabe que a veces no logran registrar a todos. Entonces, ¿cómo se hace? En ciertas localidades, se manda a hacer una nueva lista y ves cuánta gente coincide con la anterior. Cuando la coincidencia es al 100%, se dice que el censo ha registrado a todo el mundo. Pero cuando hay gente que está en la evaluación post censal que no ha sido registrada en el censo, se ve la tasa de traslape y se tiene un estimado de cuánta gente falta. Es una técnica que no es nada del otro mundo y además tiene más de cien años utilizándose para distintos problemas de este tipo.

Existen tres tipos de críticas: gente que no le gusta la cifra por motivos políticos, aunque le expliques la teoría igual va a decir que no le gusta porque la Comisión está llena de comunistas; gente que no entiende el método, porque no es fácil; y gente que cree entenderlo. Hasta ahora, yo no he leído nada escrito por alguien que haya abordado esto desde un punto de vista científico y no como un tema político.

La última persona que ha escrito algo es Silvio Rendón, profesor de una Universidad en Estados Unidos. Él, desde el inicio, ha sido muy crítico de la Comisión de la Verdad en esta cuestión, pues no cree que Sendero Luminoso

haya matado a tanta gente. Se basa en el hecho de que cuando uno junta las tres listas de la CVR, de la Defensoría y las ONGs, las personas reportadas como víctimas de parte de agentes del Estado son más que las reportadas como víctimas de Sendero. Y eso es básicamente porque, salvo la CVR, las otras dos reportaban temas del Estado. Por esto, Rendón ha escrito un texto diciendo que no se puede estimar las víctimas de Sendero usando este método. Eso es discutible porque justamente una de las personas que trabajó en la CVR, que fue consultor en este análisis estadístico, diseñó un método para resolver ese problema. Rendón cree que eso está mal y tiene su propio método. Primero reconoce que este método sí sirve y que se puede usar, porque él trata de usarlo. Hace un estimado para las víctimas del Estado y llega a la misma cifra que la CVR respecto a las provocadas por el Estado, pero dice que para Sendero no se puede estimar. Entonces, en algunas cosas estima y en algunas cosas solamente usa los conteos de la CVR. Eso, desde mi punto de vista, es hacer trampa, porque no estás estimando las víctimas de Sendero. Estás diciendo que para el Estado sí uso estadística y para el otro voy a usar simplemente conteos simples.

Ese es el trabajo más serio que busca cuestionar la estimación de la CVR y que, desde mi punto de vista, tiene serios problemas metodológicos, matemáticos, y de supuestos de partida. Rendón siempre ha dicho que no es posible que Sendero Luminoso haya matado a más gente que el Estado, porque parte del supuesto que el Estado tiene más armas, pero esa no es una razón. Hasta ahora no encuentro un cuestionamiento de orden científico, técnico y serio que me haga cambiar de opinión acerca de la confianza que tenemos.

Por otro lado, eso no quiere decir que no se pueda mejorar. Hace un par de años, un estudiante de la universidad Católica sustentó una tesis reproduciendo este ejercicio, pero con una lista que no existía en la época de la CVR, una nueva lista que compiló el Ministerio de la Mujer. Hizo una serie de cálculos y, con nueva información que la CVR no tenía en su momento, llega a las mismas conclusiones que la CVR. Si uno va al Registro Nacional de víctimas, las víctimas que están ahí reportadas son más que las que había en la Comisión de la Verdad. Eso quiere decir que hay más víctimas que hace 10 años y que haya todavía más que no conocemos y que, probablemente, nunca lleguemos a conocer.

Hasta ahora no encuentro un cuestionamiento de orden científico, técnico y serio que me haga cambiar de opinión acerca de la confianza que tenemos.

¿Qué hubiera pasado si hubieran presentado la cifra de los 24,000 muertos?

La Comisión de la Verdad pudo haber dicho: "sabemos que hay más, pero no sabemos cuántas", ahí pudo haberse quedado, pero hay técnicas que permiten estimar cuánto más hay y esta es una de ellas. Si es que se puede hacer ¿por qué no hacerlo? Eso da una imagen más verosímil de lo que pudo haber pasado. Si no lo hacía la Comisión de la Verdad, lo pudo haber hecho otra persona. ¿Por qué la comisión dio ese paso adicional? Yo creo que no hubo ningún cálculo político, yo creo que hubo mucha prudencia por parte de los comisionados. Yo fui el que propuso hacer este ejercicio, porque se ha hecho antes. La Comisión de Guatemala lo hizo, se hizo en el caso de un reporte que para Kosovo. Felizmente contamos con uno de los principales expertos en este tema.

La decisión de utilizar el resultado de esta estimación en las conclusiones fue producto de muchas reuniones. Se presentaron los resultados, el sustento empírico, el sustento teórico, vino Patrick Ball, esto fue revisado por el jefe del INEI, por estadísticos de la Universidad Católica y vino el Presidente de la Asociación Norteamericana de Estadística que revisó el trabajo y también explicó a los comisionados por qué pensaba que era un trabajo bien hecho. No fue una decisión que se tomó a la ligera. Cuando yo vi la cifra, dije que hagan el cálculo dos veces más para ver qué sale. Como todo método científico, esto es perfeccionable y si hay nueva evidencia, se puede revisar. Pero hasta ahorita no encuentro ninguna nueva evidencia y más bien, la nueva evidencia que encuentro refuerza esta conclusión.

Otro de los cuestionamientos a la CVR es que es benigna con Alan García en comparación con Fujimori.

Yo no estoy de acuerdo con esa afirmación. Durante el gobierno de Belaunde fue donde hubo mayor cantidad de víctimas y de violaciones cometidas por el Estado. Si uno ve a lo largo del tiempo, en todos los períodos, la cantidad de víctimas cometidas por Sendero fue mayor que las cometidas por el Estado. Pero el período donde el Estado cometió más violaciones fue en el período de Belaunde.

En el período de García se cometieron más violaciones que en el período de Fujimori y eso está claramente dicho en el informe. En el período de García también hay un cambio en la estrategia del Estado y hacia el final es donde el Ejército rediseña un nuevo manual de lucha contra subversiva que da más énfasis en la participación de ronda campesinas, pero también a la eliminación selectiva de los sospechosos por terrorismo que marca una pauta muy diferente al del período anterior donde era más indiscriminado.

En el período de García es también donde se da más énfasis al trabajo policial, como el GEIN. En el período de Fujimori es donde la cantidad de víctimas baja, porque ya es el momento final. Incluso si uno ve la historia del conflicto armado interno, ya a finales del gobierno de García, Sendero estaba muy golpeado en el campo, por eso tenía más acciones en la ciudad donde era más fácil capturarlo. Al final del gobierno de García empezaron a capturar cualquier cantidad de mandos senderistas y eso continuó a principios del gobierno de Fujimori. Ya hacia el final del gobierno de García, Sendero estaba en un proceso de derrota estratégica. Incluso un par de semanas antes de cambio de gobierno, casi capturan a Abimael Guzmán.

Fujimori cosecha los frutos de este cambio de estrategia, sobre todo en la policía, pero mantiene una cobertura a las acciones de eliminación selectiva. Pero a diferencia de García, Fujimori tenía una responsabilidad directa con la creación de un grupo de asesinatos selectivos que era el grupo Colina. Ese tipo de conexión no hay en el caso de García, tampoco en el caso de Belaúnde. No se encontraron indicios fuertes que conectaran una responsabilidad directa en términos penales de ordenar crímenes o que los crímenes hayan sido cometidos con conocimiento de causa. Eso no significa que ni García ni Belaunde no sabían de los crímenes, pero no es que ellos los hayan ordenado o hayan diseñado. Lo que sí hay es responsabilidad política, eso sí lo dice la Comisión de la Verdad en todas partes.

Otra cosa que también es importante es que después del autogolpe del 92 Fujimori asume directamente el control de todas las instituciones del Estado. Había más controles democráticos en los años 80 que obviamente no funcionaron, pero en el caso de Fujimori, no existía ni eso. Hay una responsabilidad directa mucho mayor.

¿Consideras que es necesario que se sigan debatiendo este tipo de cuestionamientos?

A mí me gustaría que no fuera así, pero uno tiene que defender su trabajo cuando es cuestionado. Si en algún momento aparecen errores o se evidencia que hay información nueva, también es necesario estar dispuestos a incorporar las críticas y a analizarlas de manera honesta. Yo considero que hay más información ahora que podría permitir hacer una nueva estimación. La información existe, hay datos que no había antes como los del Registro Nacional de Víctimas. Si queremos volver a hacer un estimado con fuentes más diversas, más sistemas de los que había antes, con 15 mil nuevos casos que no existían hace 10 años, yo creo que se puede hacer. Es cuestión de conseguir el tiempo y los recursos para hacer de nuevo esto y para dar

nuevas luces de si el estimado va o no va por esa magnitud. Lo que yo he observado es que la información nueva va a confirmar eso.