

Anexo: Metodología

La planificación y la ejecución de la Auditoría ciudadana supusieron enfrentar problemas inéditos y buscar soluciones innovadoras para enfrentar las exigencias de una primera evaluación sobre la calidad de la democracia costarricense. En una evaluación de este tipo ¿cómo acopiar información?, ¿cómo sistematizar y documentar esta información?, ¿cómo evaluar los estándares de la calidad de la democracia? Con el fin de abrir esta iniciativa al escrutinio público, en el presente anexo se documenta detalladamente la propuesta metodológica que orientó las respuestas a estas interrogantes en el curso de la Auditoría y se presenta las fichas técnicas de los métodos de investigación empleados. La transparencia es un deber democrático, particularmente en un ejercicio no tradicional en las ciencias políticas como este, durante el cual se desarrolló un análisis complejo y ciertamente colectivo sobre un tema poco tratado en el país: la forma en que la ciudadanía costarricense vive su democracia en una época de malestar. Por eso, a la par de la descripción sobre el proceso seguido, el lector encontrará análisis sobre las dificultades encontradas y las modificaciones al planteamiento original.

La Auditoría enfrentó las interrogantes planteadas de modo poco convencional. En esta iniciativa se combinó el examen amplio de asuntos con la exploración flexible de temas; los pasos inductivos complementaron la deducción; se puso atención a las opiniones y experiencias cotidianas de vivir en democracia; se articuló el trabajo tesonero y artesanal del equipo técnico con la contratación de decenas de investigaciones. Este abordaje poco convencional fue una respuesta a varias situaciones. En la revisión de la literatura se encontró pocos estudios empíricos sobre el tema de la calidad de la democracia y una tendencia a emplear un reducido conjunto de indicadores -basados principalmente en el juicio de expertos- que estaban poco relacionados con las

preocupaciones que dieron origen a esta iniciativa.

El anexo se centra en tres asuntos: la estrategia del trabajo de campo, los métodos de investigación y el proceso de evaluación. Como se explica más adelante, el trabajo de campo implicó el montaje de una “fábrica” de investigaciones, con la aplicación de distintos métodos de investigación. Una vez recopilada la evidencia, fue necesario organizarla en forma coherente y comprensible para los paneles de evaluadores y la posterior redacción de capítulos. En total unos tres metros lineales de documentación, debidamente ordenada en 10 archivos, uno para cada panel de evaluación, originó el posterior ejercicio de paciencia de las personas que aceptaron evaluar la calidad de la democracia costarricense.

Fue indispensable diseñar un modelo de evaluación, calibrarlo y organizar el nada sencillo trabajo de su aplicación. Sesiones de inducción, respuestas individuales de los panelistas y reuniones del conjunto de expertos, fueron elementos de una impresionante acción voluntaria de personas extraordinariamente calificadas... y también muy ocupadas. En la tercera parte de este anexo se sintetiza el modelo de evaluación aplicado.

El trabajo de campo

El trabajo de campo fue la actividad que permitió acopiar la información para la posterior evaluación de la calidad democrática. En esta sección se sintetiza la estrategia metodológica y las técnicas de investigación empleadas. Su propósito es

caracterizar las piezas que formaron parte del rompecabezas de la Auditoría.

El trabajo de campo comprendió la elaboración de 40 informes de investigación; algunos de ellos fueron resultado de estudios de alta complejidad, como la encuesta nacional, las dos rondas de grupos focales o el ejercicio de Sitios Centinela, que cuentan como 4 investigaciones dentro del total, pero que involucraron equipos de trabajo amplios y procesamientos posteriores. Otros son estudios más puntuales y menos complejos. Cabe señalar que este conteo no incluye la recopilación de registros administrativos, que implicó reiterados esfuerzos de investigación realizados directamente por el equipo técnico de la Auditoría.

La opción de enfoque: auditoría *versus* investigación académica

Una auditoría es un examen de registros e informes, realizado por especialistas distintos a los encargados de su elaboración, “para asegurar una plena rendición de cuentas y ayudar (...) al cumplimiento de las responsabilidades encomendadas”¹. Una auditoría no es cualquier tipo de evaluación, es una que tiene como condición una revisión sistemática de las cuentas (o registros), y como finalidad asegurar la rendición de cuentas por parte de quienes son responsables de ejercer las funciones auditadas, y esto de la forma más exhaustiva posible.

Los propósitos, estilos, metodologías y resultados esperados de una auditoría son distintos a los de una investigación académica. En efecto, la auditoría evalúa prácticas, pero no comprueba relaciones causales entre cosas; emplea estándares de evaluación en vez de hipótesis de investigación y su resultado esperado es la mejora de las prácticas auditadas y no el desarrollo de mejores teorías científicas, aunque puede contribuir a ello (Lijphart, 1971; Ragin, 1987; Reider, 1994; Ricchiute, 1995; Valenzuela, 1998).

La escogencia de un enfoque de auditoría en vez de una investigación académica se fundamentó en razones políticas y prácticas. Desde el punto de vista político, como se afirmó en el Prólogo, el objetivo de esta iniciativa es facilitar a las y los ciudadanos un instrumento para fortalecer la calidad de su participación cívica y su capacidad para ejercer un control democrático sobre las instituciones. El interés no era probar una hipótesis que explicara las razones por las cuales un determinado desempeño democrático ocurre, sino documentar exhaustivamente dicho desempeño. Desde el punto de vista práctico, se trataba de una primera exploración sobre un tema poco estudiado y los estándares de evaluación establecidos por el Consejo Consultivo abarcaban un conjunto amplio de temas. Como toda exploración, debía existir flexibilidad, dentro de los temas tratados por los estándares, para investigar asuntos que en principio

CUADRO 1A.1

Diferencias entre una auditoría y una investigación académica

Criterios de comparación	Auditoría	Investigación académica
Propósito	Evaluar prácticas	Explicar relaciones entre objetos
Estilo ^{a/}	Comprensivo, no selectivo	Selectivo
Participación	Se requiere la participación de las personas auditadas durante la auditoría, para discutir los resultados de la evaluación	No se requiere la participación de las personas bajo estudio para discutir los resultados de la investigación
Formulación de hipótesis	No	Sí
Resultado esperado	Mejora de prácticas auditadas	Desarrollo de mejores teorías
Otros resultados (prioridad secundaria)	Desarrollo de mejores teorías	Varios, de acuerdo con los intereses del investigador

a/ Comprensivo significa que una auditoría cubre todos aquellos asuntos que se considere relevantes para los propósitos de la evaluación. Selectivo significa que una investigación académica procura delimitar al máximo el número de asuntos bajo estudio, a efecto de ganar la mayor profundidad posible.

FUENTE: GUTIÉRREZ Y VARGAS-CULLELL, 1998.

podían ser relevantes y sobre los cuales no había razones conceptuales o metodológicas para excluirlos *a priori*.

La estrategia de investigación

La decisión de realizar una auditoría y no una investigación académica influyó profundamente sobre el trabajo de campo. En efecto, su misión era recopilar la mayor cantidad posible de información sobre lo que estaba ocurriendo en los temas tratados por los estándares democráticos establecidos por el Consejo Consultivo, con el fin de permitir, posteriormente, una primera evaluación de su cumplimiento. Ello planteó la necesidad de adoptar, como punto de partida, una estrategia de investigación abierta a distintos métodos y perspectivas de análisis.

La estrategia de investigación en el trabajo de campo se sustentó en cinco premisas:

- Fomentar la más amplia participación social, académica e institucional en el proceso de investigación.
- Estimular perspectivas multidisciplinarias en la investigación de los estándares de calidad de la democracia.
- Trabajar los temas como módulos independientes.
- Combinar los métodos de investigación en el estudio de un asunto.
- Aplicar una diversificación metodológica.

Primera premisa: Participación social, académica e institucional

El trabajo de campo activó un entramado de actores sociales, políticos, institucionales y académicos que debían producir insumos para una obra colectiva con valor intelectual y legitimidad suficientes para adentrarse en la evaluación de la calidad de la democracia costarricense. La participación de estos actores en el proceso de investigación se propició por diversos medios, como la convocatoria a líderes sociales y políticos para que participaran en la aplicación de ciertos métodos de investigación (por ejemplo, grupos focales), la solicitud de opiniones legales e investigaciones cortas a decenas de investigadores y analistas y la ayuda para que las y los investigadores de la Auditoría tuviesen acceso a estudios y registros administrativos poco conocidos o sistematizados.

Segunda premisa: Multidisciplinariedad

El trabajo de campo articuló una red académica y técnica compuesta por centros de investigación de las universidades públicas, funcionarios vinculados a los poderes del Estado y diversas entidades de la administración pública, organizaciones privadas de reconocida trayectoria y experiencia en el campo de la investigación social, así como un conjunto de actores sociales y políticos de la más variada procedencia e incidencia en la vida nacional. Esta red permitió combinar investigadores y analistas de disciplinas vinculadas de una manera u otra con el estudio de la calidad de la democracia -ciencias políticas, sociología, administración pública, antropología, derecho y estadística- con la experiencia práctica del funcionamiento de la institucionalidad costarricense y, en general, con la manera en que las y los ciudadanos viven la democracia en los ámbitos de la relación con el poder público y con sus conciudadanos.

Tercera premisa: Módulos independientes

En el trabajo de campo se aplicó simultáneamente varios métodos de investigación. Por tratar cada uno temáticas distintas, los métodos fueron empleados por diferentes equipos de investigadores. Cada método funcionó como un módulo independiente de la Auditoría². Esta estructura modular funcionó también para el tratamiento de los estándares y los indicadores³.

Sin embargo, como se verá más adelante, se propició ciertos encadenamientos y correlaciones entre métodos, con el fin de verificar resultados o complementarlos con nueva información. Por ejemplo, la información recabada por un método (v.gr. análisis jurídico o recopilación de registros administrativos) enriqueció el trabajo de los investigadores que estaban aplicando otro método (v.gr. una ponencia especializada). Empero, no se aplicó una regla general para los encadenamientos, con excepción del pareo entre encuesta y grupos focales y entre todos los métodos de investigación y los paneles de expertos, desarrollados de manera exclusiva y dedicada hacia el final del trabajo de campo. Cuando fue necesario encadenar esfuerzos, ello se señaló en el protocolo específico de investigación para cada indicador de los estándares⁴.

Cuarta premisa: Combinaciones metodológicas

Los estándares de alta calidad democrática -y la mayoría de sus indicadores- fueron abordados, como regla general, por más de un método de investigación. La combinación de métodos, propia de las auditorías, procuró cumplir con tres objetivos. Primero, decantar las fortalezas y debilidades de

cada método de investigación en el estudio de las prácticas democráticas. Segundo, recopilar distintas perspectivas sobre un mismo tema, con el fin de enriquecer la valoración sobre el cumplimiento de los estándares (fase de evaluación). Finalmente, contrastar (o en su defecto complementar) la información recabada mediante la aplicación de un método, con la recopilada por otros, para identificar, con miras a futuros ejercicios, las combinaciones metodológicas con mayor potencial y descartar aquellas con menor potencial⁵.

La combinación metodológica en el trabajo de campo se guió por los siguientes criterios:

- Los métodos cuantitativos, como la encuesta y los Sitios Centinela, fueron usualmente acompañados por métodos cualitativos, como los ejercicios experimentales. La combinación característica fue, por ejemplo, la encuesta con un ejercicio experimental o una ponencia especializada.
- Cuando no fue posible la anterior combinación cuantitativa y cualitativa, se realizó una combinación entre métodos cualitativos.
- Muy excepcionalmente, un indicador fue investigado por un solo método cualitativo.
- Casi todos los estándares (con excepción de uno) fueron abordados mediante una combinación

de registros administrativos y recopilación bibliográfica o de estudios previos.

Quinta premisa: Diversificación metodológica

El trabajo de campo incluyó, como se presenta más adelante, nueve métodos de investigación y dos tipos de acciones de recopilación de información. Se trató de auscultar la calidad de la democracia costarricense empleando la mayor cantidad posible de métodos (una calificada excepción a esta premisa fue la verificación del cumplimiento de los estándares del Umbral Mínimo de Garantías Democráticas (UMGD), concentrado en un análisis especializado y presentado en el capítulo 2). Los métodos de investigación se ordenaron en cuatro categorías: métodos de consulta, métodos experimentales, métodos interpretativos y acciones de recolección documental (cuadro 1A.2).

Los métodos de consulta fueron aquellos mediante los cuales se auscultó la opinión de la ciudadanía costarricense, ya fuera mediante procedimientos de muestreo estadístico, o a través de procedimientos estadísticamente no representativos. Se empleó cuatro métodos de consulta: una encuesta nacional, los grupos focales, los Sitios Centinela y el juicio grupal ponderado.

Los ejercicios de observación experimental consisten en aproximaciones directas al comportamiento de actores en micro-situaciones reales, para analizar prácticas relevantes para la calidad de la democracia. En la Auditoría, cada ejercicio contó

CUADRO 1A.2

Métodos de la Auditoría

Tipo	Método
Métodos de consulta	Encuesta nacional Grupos focales Sitios Centinela Juicio grupal ponderado
Métodos experimentales	Ejercicios etnográficos de observación experimental Observación en medios de comunicación colectiva
Métodos interpretativos	Ponencias especializadas Investigaciones temáticas Análisis jurídicos
Recolección documental	Recopilación y sistematización de registros administrativos Recopilación bibliográfica y de estudios previos

con su correspondiente definición de objetivos, protocolo de observación y criterios para la sistematización, tabulación y presentación de resultados. Se aplicó dos tipos de ejercicios: etnográficos y de observación de medios de comunicación colectiva.

Una mirada retrospectiva sobre la estrategia del trabajo de campo

La realización de un trabajo de campo de esta complejidad tuvo, en la práctica, ventajas y desventajas. Ciertamente, la aplicación de diversos métodos permitió recopilar un cúmulo de información valiosa sobre el funcionamiento de la democracia costarricense, mucha de ella inédita. Permitted, también, valorar la idoneidad de los métodos para los fines propuestos. Por ejemplo, el método de los Sitios Centinela (véase más adelante) ofreció información novedosa y exhaustiva. Convocó a una red de investigadores de diversas disciplinas y corrientes de pensamiento al estudio sobre la calidad de la democracia. En materia de los estudios políticos, no existen antecedentes en Costa Rica sobre una investigación de esta naturaleza.

Sin embargo, es preciso reconocer también las desventajas del planteamiento participativo. En primer lugar, fue difícil lograr que los plazos previstos para el trabajo de campo se cumplieren. La administración de una red tan amplia de investigadores fue una tarea singularmente compleja, dados los recursos con que contaba el equipo técnico. La mayor parte de las investigaciones fue entregada con uno o dos meses de retraso, y los últimos informes fueron recibidos seis meses después de lo previsto⁶. En segundo lugar, buena parte de esta red de investigadores funcionó *ad honorem* o con un pago simbólico, más bien por el convencimiento de las personas sobre la utilidad de la Auditoría. Aunque ello atestigua la importancia de esta iniciativa, indudablemente dificultó, en algunos casos, el control de calidad sobre las investigaciones.

Una mención aparte merece la extensión de los plazos reales de investigación, que restó tiempo para las etapas subsiguientes de la Auditoría, especialmente el procesamiento de la información recopilada y la especificación del modelo de evaluación (este último asunto será tratado más adelante). La reducción del tiempo disponible para el procesamiento complicó lo que ya era, debido a la naturaleza del trabajo de campo, una tarea difícil: la clasificación de la información por estándar, pregunta de evaluación e indicador. Fueron necesarias sesiones maratónicas de trabajo, durante

varias semanas, por un equipo dedicado exclusivamente a la clasificación temática de la información, al traslado de esa información a un mismo procesador de palabras (Word 97) y hojas electrónicas (Excel 97) y la creación de formatos uniformes para los textos, cuadros y gráficos. Pese al enorme y cuidadoso esfuerzo, en ciertos casos la clasificación no fue adecuada -asunto que los evaluadores hicieron notar-, o hubo reiteraciones de información. Es preciso reconocer que se había subestimado la dificultad de esta tarea.

El acortamiento del tiempo disponible impidió realizar un estudio acerca de la consistencia técnica de los métodos empleados, particularmente los de consulta y los experimentales. Con la ayuda de un miembro del Consejo Consultivo, esta labor se hizo en el caso de la encuesta nacional (Gómez, 2000). Ello permitió detectar problemas en la formulación de ciertas preguntas, cuyos resultados no fueron incluidos en los capítulos respectivos. Asimismo, la información proveniente de los registros administrativos fue minuciosamente verificada. Sin embargo, esta labor está pendiente para el resto de los métodos. Finalmente, no hubo oportunidad para reprocesar la información estadística de los Sitios Centinela, más allá de los cuadros generales y los cruces de variables originalmente solicitados.

Aún así, el balance es positivo para los fines propuestos. Se demostró que era posible diseñar y ejecutar un proceso técnico y socialmente complejo con un equipo que contaba con pocos recursos humanos y financieros. Las investigaciones “descubrieron” temas y hechos poco conocidos y debatidos en Costa Rica, como el de las diferencias en la calidad de los gobiernos locales o el trato de las instituciones públicas a las personas. La información recopilada proporcionó en general el material que las y los evaluadores requerían para sus juicios. Fue, en síntesis, un trabajo de campo cuyas lecciones sientan una valiosa pauta para futuras evaluaciones sobre la calidad de la democracia en Costa Rica.

Fichas técnicas de los métodos de investigación

Métodos de consulta

Ficha 1: Encuesta nacional



Descripción:

Encuesta domiciliar nacional, seleccionando zona específica de estudio, con el fin de establecer comparaciones entre regiones. En las zonas seleccionadas se reforzó la cantidad de entrevistas por

realizar, para poder hacer inferencias y comparaciones entre regiones. Durante el procesamiento final, se utilizó factores de ponderación para adecuar el tamaño de las submuestras regionales a la proporción real de la población.

Características:

La población de interés estuvo constituida por ciudadanos y ciudadanas costarricenses de entre 18 y 64 años de edad. La cobertura fue tanto rural como urbana. Las zonas geográficas seleccionadas fueron: Área Metropolitana; resto del valle urbano; resto del valle rural; Zona Norte; Zona Atlántica típica; resto zona atlántica; Zona Sur; Zona Pacífica.

La información fue recolectada en entrevistas domiciliarias directas con las personas seleccionadas, utilizando un cuestionario estructurado y altamente precodificado, con más de 50 preguntas, abiertas y cerradas. La duración total de cada entrevista fue de aproximadamente 45 minutos.

La muestra fue de 1.618 sujetos, lo que permitió trabajar con un nivel de confianza del 95% y un error máximo de 2,4 puntos porcentuales. El muestreo fue probabilístico, polietápico y utilizó la información censal básica y la cartografía más reciente del país. La unidad primaria del muestreo fueron los segmentos censales, con probabilidad proporcional a su tamaño: 800 para la Gran Área Metropolitana (GAM) y 160 para cada una de las otras cinco regiones. En cada segmento censal seleccionado se visitó 10 viviendas, escogidas sistemáticamente a partir de un arranque aleatorio. Las personas por entrevistar se seleccionaron de acuerdo con el cumplimiento de criterios y cuotas de acuerdo con características sociodemográficas. Si la persona no se encontraba, se sustituía visitando la primera casa vecina. La encuesta se realizó entre el 23 de abril y el 12 de mayo de 1999, por 15 encuestadores y 4 supervisores debidamente adiestrados, en un horario de 2 p.m. a 8 p.m. durante los días entre semana, y desde la mañana los sábados y domingos.

Procesamientos posteriores para su uso:

Se reprocesó las preguntas relacionadas con calidad de la democracia, para determinar si existen diversos modelos mentales de democracia y si éstos tienen correspondencia con las características de las personas entrevistadas (ingreso, región, educación, etc.). Mediante un análisis factorial se construyó escalas (variable complejas) sobre diversos temas.

Principales adaptaciones del método para la Auditoría:

La necesidad de comparación entre regiones supuso una ampliación del número de entrevistas más allá de la proporción según su tamaño (diseño desproporcionado), asunto que fue revertido al momento de procesar los datos generales.

Créditos:

Empresa UNIMER. Directora de proyecto: Valeria Lentini. Asesor específico para la encuesta y el muestreo: Miguel Gómez. Los reprocesamientos posteriores estuvieron bajo la responsabilidad de Elisa Sánchez y Miguel Gómez, con la asistencia de Mauricio López.

Principales estándares sobre los que aporta información:

Estándares asociados al acceso a información veraz, oportuna y objetiva para evaluar las candidaturas; prejuicios sociales y culturales que enfrentan los candidatos y candidatas; exigencia de rendición de cuentas; creencia en la democracia; gobiernos locales; expresión libre y oportuna sobre asuntos públicos; medios de comunicación; escrutinio público de las organizaciones de la sociedad civil; conocimiento de los derechos y deberes; aplicación en asuntos públicos de normas democráticas; selección de servidores públicos; trato a la ciudadanía por parte de las instituciones; reglas de acceso a servicios públicos; ejercicio de la tolerancia; unidad cívica nacional y conocimientos cívicos; protección de derechos de grupos étnicos y de género; vida en el barrio.

Hallazgos de mayor interés:

Los principales aportes específicos se refieren al concepto y desagregación de la calidad de la democracia, el maltrato al ciudadano y la vida cotidiana de la gente en el barrio, sus organizaciones, etc.

También resultan de particular interés las relaciones que se logró establecer entre el concepto de calidad de la democracia y las características de la población. Por ejemplo, a menor ingreso, la noción de calidad de la democracia está más asociada a las cualidades del liderazgo político y a ciertos resultados concretos en materia de bienestar social y políticas públicas. A mayor ingreso, la calidad de la democracia está más asociada a la calidad de las reglas del juego político.

Documentos disponibles en la página web:

Formulario de encuesta e informe final con los resultados.

Descripción:

Los grupos focales son un recurso -complementario o no a una encuesta- que permite organizar sesiones con pequeños grupos de ciudadanos representativos del conglomerado social, o de un segmento de éste, a efecto de profundizar en la indagación o interpretación de ciertos temas. En general, el interés radica en aplicar el instrumento a personas seleccionadas con base en criterios previamente definidos. En algunos casos, el grupo focal puede realizarse con personas que tengan conocimientos especializados (por razón de su formación o por su experiencia de vida).

Características:

Se realizaron 6 sesiones temáticas y 6 sesiones en regiones geográficas.

Sesiones temáticas:

Estas sesiones tuvieron como propósito determinar la percepción de determinados grupos de ciudadanos sobre aspectos esenciales del sistema político costarricense. Las 6 sesiones se realizaron entre el 7 y el 16 de setiembre, con la participación de líderes políticos, sociales y académicos del país, y con una duración promedio de dos horas y media cada una.

En total participaron 37 personas (12 mujeres y 25 hombres): diputados (PUSC, PLN, Fuerza Democrática), exdiputados, excandidatos a la Presidencia, exministros, funcionarios públicos de alto nivel jerárquico, abogados (autoridades de la Corte Suprema de Justicia, expresidentes del Colegio de Abogados), representantes de organizaciones de mujeres, dirigentes indígenas, antropólogos, sociólogos, politólogos, presidentes de partidos políticos, la Segunda Vicepresidenta de la República y otros⁷. En todas las sesiones hubo participantes de ambos géneros, excepto en la que trató el tema de la igualdad real de la mujer, en la cual sólo participaron mujeres.

De manera natural y espontánea, los participantes en las sesiones establecieron interrelaciones entre la mayoría de los temas. Eso se dio con más intensidad en el tema del “desencanto con la política”, presente en todas las sesiones.

Sesiones geográficas:

Se realizaron 6 sesiones de grupo para indagar entre ciudadanos y ciudadanas cómo viven los costarricenses la democracia en su experiencia cotidiana. Se exploró la convivencia en las comuni-

dades, con las organizaciones y líderes locales, con la municipalidad, etc.

Las sesiones estuvieron distribuidas de la siguiente manera: hombres y mujeres, en igual proporción, de 25 a 45 años, de cualquier profesión, de nivel socioeconómico medio-amplio, vecinos de San José y Cartago; Heredia y Alajuela; Pérez Zeledón; Zona Norte (San Carlos); Región Pacífico; Región Atlántica.

Créditos:

Empresa UNIMER. Directora de proyecto: Alejandra Mesén.

Principales estándares sobre los que aporta información:

Cada sesión temática centró su discusión en uno de los siguientes temas, orientados por la respectiva guía: “Desencanto con la política”, “Control ciudadano de la gestión pública”, “Los grupos étnicos como parte de la unidad cívica nacional”, “Igualdad real de la mujer”, “Las contralorías de servicios públicos” y “El funcionamiento del Poder Judicial”. Los participantes en las sesiones sobre “Control ciudadano de la gestión pública” y “Funcionamiento del Poder Judicial”, espontáneamente unieron sus temas con el de desencanto político, coincidiendo en la forma en que explican este último fenómeno social, así como en sus implicaciones para el país y sus instituciones. Los grupos focales que analizaron el tema de los “Grupos étnicos como parte de la unidad cívica nacional” y el de “Igualdad Real de la Mujer” coincidieron, en general, con los otros tres citados, en lo relativo a la existencia de un desencanto político.

Los grupos focales geográficos recabaron información sobre una importante variedad de asuntos relacionados con los temas de la cultura cívica y la participación ciudadana en las organizaciones de la sociedad civil y en las políticas públicas. Permitieron establecer similitudes y diferencias locales.

Hallazgos de mayor interés:

Se recabó información valiosa sobre la discriminación a nicaragüenses, las debilidades en las prácticas de rendición de cuentas en las instituciones nacionales y locales y percepciones sobre el funcionamiento del Poder Judicial.

Documentos disponibles en la página web:

Informe final de las Sesiones Temáticas.
Informe final de las Sesiones Geográficas.

El método de los Sitios Centinela incluye la aplicación de una encuesta de hogares, entrevista a informantes clave, realización de un grupo focal y de un juicio grupal ponderado en cantones seleccionados según los resultados de un juicio grupal ponderado con miembros del equipo técnico de la Auditoría y expertos invitados.

Descripción, características e instrumentos:

Los Sitios Centinela constituyen un método de investigación aplicable en un panel de lugares seleccionados (microuniversos) según criterios y por razones precisas, de forma sistemática e informada, mediante un juicio grupal ponderado con expertos. Se basa en el levantamiento de información cuantitativa y cualitativa con instrumentos diversos, tales como encuesta aleatoria en grupos de hogares, entrevistas a informantes clave y grupos focales. Permite aproximar conclusiones razonablemente objetivas y generalizables sobre temas específicos. No es un método basado en la inferencia estadística. Fue inicialmente empleado en epidemiología para el estudio rápido y efectivo de causas o condiciones de epidemias en sitios determinados, con el fin de identificar y anticipar tendencias más generales.

La Auditoría empleó esta metodología para valorar el dominio de los gobiernos locales, con un único estándar (los gobiernos locales gobiernan democráticamente) y el correspondiente conjunto de indicadores y variables (elecciones libres y limpias; escrutinio público y rendición de cuentas; participación ciudadana).

Se aplicó en 7 cantones (Puntarenas, Carrillo, Belén, Talamanca, Alajuelita, Palmare, San Pedro de Montes de Oca), seleccionados mediante criterios elaborados en un juicio grupal ponderado con expertos y miembros del equipo técnico. Los criterios de selección fueron: variación del grado de participación electoral cantonal, diferentes experiencias municipales reconocidas, diversidad de estratos sociales, niveles de desarrollo humano, niveles de organización social, gasto municipal per cápita, deserción escolar, carácter rural o urbano, rasgos étnicos distintivos y principal actividad económica. De ese conjunto de criterios se presenta, en un anexo adjunto, algunos datos cuantificables de gran relevancia.

Se visitó 3.200 hogares, se entrevistó a 141 personas entre informantes clave y autoridades municipales y se realizaron 7 grupos focales y 7 juicios grupales ponderados (uno por municipio).

Principales adaptaciones del método para la Auditoría:

La adaptación principal fue sobre los temas y elementos por observar. Un elemento adicional de importancia en este método es que supone una devolución de la información a la comunidad estudiada.

Créditos:

El juicio grupal ponderado para seleccionar los municipios por estudiar y el trabajo en los cantones seleccionados fue realizado por William Brenes y René Martorell.

Principales estándares sobre los que aporta información:

Aportó información sobre el dominio y el estándar sobre gobiernos locales, trato al ciudadano, participación ciudadana y procesos electorales.

Idoneidad del método y complementación necesaria:

Los resultados obtenidos por la combinación de instrumentos de los Sitios Centinela son de gran valor, tanto así, que una de las conclusiones metodológicas de la Auditoría se refiere a la conveniencia y la necesidad de utilizar el método más allá del estudio de los estándares asociados a los gobiernos locales.

Hallazgos de mayor interés:

El contraste entre las percepciones de la población y las “autopercepciones” de los funcionarios y representantes a nivel local. La existencia del “clientelismo” y su relación con la calidad de la democracia en los municipios.

Disponible en página web:

Informe final sobre Sitios Centinela.

Ficha 4: Juicio grupal ponderado



Descripción:

El juicio grupal ponderado es un método para la evaluación cualitativa y la toma de decisiones, aplicable a programas que enfrentan problemas calificados como no específicos por las disciplinas relevantes. Fue originalmente diseñado para temas de salud, como esquema alternativo a los juicios clínicos de especialistas. Resulta, en tal situación, un método mejor que la conjetura de expertos, pues llega a decisiones más consistentes. Se emplea para: valorar el éxito de un proyecto existente; estimular el potencial de éxito de un programa existente; desarrollar las prioridades

que deben ser más consideradas, por los planificadores, en la formulación de un proyecto. Puede ser aplicado en procesos de evaluación de proyectos en los que las variables por considerar son independientes una de la otra y pueden ser sopesadas en tal forma que exista una relación positiva entre sus valores y el éxito del proyecto. Por ser un ejercicio usualmente realizado bajo condiciones de escasa o nula información estadística previa, las variables dependientes (los problemas que están siendo tratados) pueden no ser excluyentes entre sí y mostrar cierto grado de traslape; la selección de las variables independientes (componentes del programa) no obedece a una teoría o hipótesis previamente establecida, sino a criterios de expertos sobre los asuntos relevantes para enfrentar ciertos problemas, y las prioridades resultantes de la deliberación del grupo podrían no corresponder con los resultados de otros estudios.

En un juicio grupal ponderado los pasos son: a) identificación de las personas u organizaciones relevantes que son actores en el programa; b) reunión del panel con un moderador; c) se fija el tema, mediante una pregunta generadora; d) cada miembro del panel responde a la pregunta en forma escueta, sin argumentación ni discusión, la que se anota a vista de todos; e) una vez recopiladas las respuestas, se someten a discusión una a una, brevemente; f) mediante votación, se reduce el número de respuestas; g) se procede a aclarar las respuestas anotadas, mediante la discusión de los patrones de votación inconsistente; cuando corresponda se pueden eliminar duplicaciones o lograr una mejor clasificación; h) se valoran y seleccionan respuestas mediante la asignación de prioridades, para quedarse con unas 10 respuestas; i) se levanta la información sobre las percepciones de éxito de las respuestas, mediante una valoración de los miembros del panel; j) en pasos sucesivos, y con cálculos específicos, se llega a la selección de un conjunto pequeño de asuntos sobre los que se expresa acuerdo en cuanto a su importancia y prioridad. El sentido de este ejercicio es generar un proceso social tendiente a extraer recomendaciones y preferencias, reducir esas preferencias a unas pocas esenciales (priorizar), producir una estadística simple y significativa sobre esas prioridades y aislar las fuentes de error (Valdez, 1985).

Características:

El juicio grupal ponderado fue aplicado como uno de los métodos de los Sitios Centinela, como se indicó en la ficha 3 y, además, como una técnica para la exploración de la calidad de la vida

democrática dentro de las organizaciones de la sociedad civil. Para tales efectos, se realizó un panel con 15 participantes de organizaciones sociales (sindicatos, organizaciones de capacitación, asociaciones y federaciones sociales y ambientalistas).

Créditos:

René Martorell y William Brenes se encargaron de dirigir el juicio grupal ponderado.

Principales estándares con los que se relaciona:

La vida democrática en las organizaciones de la sociedad civil. El propósito central del juicio grupal ponderado fue identificar y poner de relieve las variables más significativas en las siguientes asuntos: aspectos formales: régimen electoral al interior de las organizaciones de la sociedad civil; rendición de cuentas: formas que asume, frecuencias de uso; representatividad: relación de la dirigencia con las bases; información a las bases: cómo se informa, periodicidad; protección de derechos: libertades de los agremiados; la “línea” de la organización: cómo se construye; y vida democrática en las organizaciones de la sociedad civil.

Hallazgos de mayor interés:

Fue muy significativa la valoración sobre la vida democrática en la sociedad civil.

Disponible en página web:

Informe final de Juicio Grupal Ponderado sobre las organizaciones de la sociedad civil.

Métodos experimentales

Ficha 5: Ejercicios etnográficos de observación experimental



Descripción:

La etnografía usa las clásicas técnicas de acopio de información por medio de la observación (observación participante, conversación, entrevista no estructurada, fotografía y otros medios gráficos, análisis de contenido de documentos, de las historia de vida y de los testimonios). Es una metodología estrictamente cualitativa, aunque no desdeña los datos cuantitativos que aparecen. La etnografía es un ejercicio de traducción de cultura; busca dar voz directa a la población en estudio.

Características:

Mediante los ejercicios etnográficos se procuró que la ciudadanía tuviera voz directa para manifestar sus percepciones y opiniones en torno a la

democracia y su ejercicio en Costa Rica. Los experimentos aportaron información sobre 12 de las 33 aspiraciones identificadas, en algunos casos mediante varios experimentos por estándar. El fin específico fue combinar los resultados de los experimentos con los resultados de otros métodos, al apreciar las opiniones, percepciones, valores y prácticas, mitos (en el sentido antropológico de explicación de origen de un código cultural) y expectativas sobre la democracia en Costa Rica que tienen la ciudadanía y otros sectores de población que viven y trabajan en el país.

En el transcurso de los estudios se tomó nota detallada de los experimentos mediante bitácoras, grabaciones y apuntes, en su modalidad de registro de expresiones individuales o grupales. Cada experimento se realizó con varios grupos y en lugares distintos del territorio nacional, para captar diferencias y tender a cierta representatividad, sin pretender alguna representatividad estadística.

Para cada experimento se produjo un documento con la descripción del ejercicio y su correspondiente análisis, con anexos de documentos o entrevistas probatorias de los casos más excepcionales. Es importante considerar que este es un método que permite ilustrar percepciones y prácticas políticas, pero sus conclusiones no son generalizables a la población del país o de una región.

Créditos:

El antropólogo Carlos Borge planeó y desarrolló los ejercicios de observación.

Principales estándares sobre los que aporta información:

Se realizaron experimentos en 12 estándares, referidos a prejuicios sociales y culturales que disminuyen la elegibilidad de candidaturas; escrutinio público del financiamiento de partidos; reglas democráticas en la vida de las organizaciones sociales; protección de los derechos y acceso a la justicia; libertad de expresión medios de comunicación y opinión pública; adquisición de ciudadanía; creencia en valores democráticos; trato degradante al ciudadano por parte de las instituciones; identificación con una unidad cívica nacional.

Algunos ejemplos de observaciones experimentales:

- Se muestra fotos uniformes de personas de distinto fenotipo (“cartagos”, nicaragüenses, negros, indios, chinos y europeos nórdicos), sexo y ubicación social, a grupos de personas de distintas partes del país y se les consulta por quién

votarían y las razones de su decisión.

- Ante los mismos grupos se da lectura a un pequeño texto con una misma declaración política (neutral, no tendenciosa y positiva) y se pone como autor a un indio (Benjamín Mayorga), un nicaragüense (Pablo Antonio Cuadra), un político del Valle Central (Luis Fishman y Walter Muñoz) y una mujer (Margarita Penón), para escuchar y observar las reacciones de distintos grupos del país. La cartulina con el texto lleva una foto de esas personas y se le entrega una distinta a cada persona del grupo.
- Se identifica la proclividad de un medio de comunicación por un tema nacional (la línea editorial denota las simpatías de un medio sobre determinado tema) y se manda artículos y cartas a favor y en contra de dicha línea política, para registrar si se da espacio o no a estas posiciones.
- Se escribe y publica un artículo cautamente provocador sobre derechos de los negros o de los indios a una cuota de representación electoral, sobre el derecho a ser electo en un cargo público de una persona que tenga un origen claramente nicaragüense, o sobre un asunto ecológico, para registrar las reacciones del público.

Disponible en página web:

Informe final de los ejercicios de observación experimental.

Ficha 6: Observación experimental en los medios de comunicación colectiva



Descripción:

Mediante la observación no participativa y experimental en dos medios de comunicación colectiva se acopió información para el capítulo sobre la calidad democrática de la opinión pública.

Características:

Durante una semana se llevó un registro sistemático y minucioso de lo que sucede en la redacción y en el equipo periodístico de un medio de comunicación. Aceptaron participar en el ejercicio los diarios *La Nación* y *Al Día*. Se observó específicamente la participación de la ciudadanía en el proceso de recolección de la información y el seguimiento que los medios brindan a las noticias; la calidad de las respuestas que las fuentes informativas brindan a los periodistas (transparencia), particularmente en el

caso de las fuentes nacionales; el ambiente de trabajo en los medios de comunicación, incluyendo la iniciativa periodística, la eficacia para el abordaje, cómo se seleccionan temas para elaborar una noticia, etc.

Para esto se elaboró un protocolo para la observación y criterios para la sistematización, tabulación y presentación de los datos recabados. Se trabajó con bitácoras minuciosas (tiempo, lugar, circunstancias, perfil del personaje y otros hechos y detalles que sucedan). Para cada ejercicio se preparó un informe final con el análisis general de los resultados, el estándar y el indicador considerados, descripción y análisis de lo observado, anexos con documentos relevantes, entrevistas y otros materiales complementarios, así como una valoración general sobre cada indicador y aspiración evaluados.

Los principales instrumentos para el análisis fueron: bitácora general; mapeo del funcionamiento normal de la redacción; entrevista con los directores de medios; entrevista privada con cada uno de los periodistas; bitácora especializada de asistencia a reunión de editores; formularios que los periodistas debían llenar diariamente; hoja de tips; hoja de fuentes de información.

Créditos:

Marjorie Ross se encargó del diseño de instrumentos y de realizar las observaciones en los dos medios de comunicación colectiva.

Principales estándares sobre los que aporta información:

Escrutinio de los asuntos públicos por los medios de comunicación colectiva.

Hallazgos de mayor interés:

Se pudo recabar información precisa sobre la participación de las personas en la elaboración de las noticias, las limitaciones del marco legal para el ejercicio del periodismo y la relación entre medios, sectores políticos y empresas publicitarias.

Disponible en página web:

Informe final de Observación Experimental en medios de comunicación colectiva.

Métodos interpretativos

Ficha 7:

Ponencias especializadas



Descripción:

La Auditoría contó con 10 ponencias especializadas. Las ponencias son estudios sobre temas puntuales, la mayoría abordados con una perspectiva

CUADRO 1A.3

Ponencias especializadas contratadas

Autor	Tema
Víctor Ramírez	Empleo de prejuicios en campañas electorales.
Jorge Quesada	Fuentes información política.
Saúl Weisleder	Financiamiento de los partidos políticos.
Fabián Volio	Medios de protección de derechos civiles y políticos.
Ivonne Jiménez	Medios comunicación colectiva.
Ronny Rodríguez	Funcionamiento democrático del Poder Legislativo.
Manrique Jiménez	Competencias de poderes y conflictos interorgánicos.
Roberto Vargas	Ejercicio del periodismo y libertad de expresión.
Epsy Campbell	Protección equitativa y eficaz de los derechos de grupos étnicos y de género.
Laura Arguedas	Estudio sobre el funcionamiento de las organizaciones de la sociedad civil.

disciplinaria no jurídica (por ejemplo, política, sociológica, estadística o de psicología social). Combinan el análisis con la recopilación de información de fuentes primarias y secundarias de investigación.

Características:

Responden a la idea de asignar la tarea de investigación y análisis a una persona, que tiene amplia libertad para desarrollar el trabajo, incluyendo la red de colaboradores. En 4 casos las ponencias trataron sobre un solo indicador; las 6 restantes investigaron temas relacionados con varios indicadores (entre 3 y 7 indicadores).

Principales estándares sobre los que aporta información:

Las 10 ponencias especializadas se relacionan con 33 indicadores, asociados a 10 estándares.

Hallazgos de mayor interés:

Mediante algunas ponencias fue posible estudiar el funcionamiento de la Asamblea Legislativa y el de los medios de comunicación colectiva en Costa Rica. En este último tema resulta interesante la información sobre la relación entre medios de comunicación y agencias publicitarias.

Documentos disponibles en página web:

Todas las ponencias especializadas.

Ficha 8:

Investigaciones temáticas



Descripción:

La Auditoría realizó 4 investigaciones temáticas. Éstas son estudios complejos y de mayor

profundidad que las ponencias especializadas, sobre varios estándares e indicadores previamente agrupados en temas generales. Estos trabajos implicaron la recopilación y procesamiento de una diversidad de fuentes de información secundaria (encuestas de opinión, registros administrativos) y terciaria (otras investigaciones).

Créditos:

Según se resume en cuadro anterior.

Principales estándares sobre los que aporta información:

Las investigaciones temáticas aportaron información novedosa, sustancial y crítica sobre las buenas y malas prácticas del ejercicio democrático en los temas seleccionados. Los temas son los siguientes:

1. “Rendición de cuentas y representación responsable”. Trata 15 indicadores relacionados con temas clave en el ámbito de la representación con rendición de cuentas.
2. “Democracia local”. Trata 14 indicadores relacionados con la vivencia democrática y la representación en el ámbito municipal.
3. “Trato ciudadano”. Trata 20 indicadores relacionados con la civilidad y el pluralismo político.
4. “Identidad común con la nacionalidad costarricense (unidad cívica nacional)”. Trata 10 indicadores relacionados con la tolerancia, el pluralismo y la pertenencia a la nacionalidad costarricense.

En total, 37 indicadores correspondientes a 16 estándares fueron abordados por alguna de las investigaciones temáticas.

Hallazgos de mayor interés:

Vía ejemplificación puede destacarse el tipo de aporte de las investigaciones temáticas: por una

CUADRO 1A.4

Investigaciones temáticas contratadas

Autor	Tema
Hugo A. Muñoz	Rendición de cuentas.
Roy Rivera	Democracia a nivel local.
Max Esquivel y Juany Guzmán	Trato de las instituciones al ciudadano.
Rodolfo Cerdas	Unidad cívica nacional.

parte, la distinción entre una unidad nacional incluyente, originada en valores cívicos, y otra, de carácter excluyente, basada en una supuesta homogeneidad étnica; por otra parte, el trato al ciudadano dispensado en las instituciones públicas como un tema propio de la calidad de la democracia, abordado mediante el estudio de quejas presentadas ante la Defensoría de los Habitantes de la República.

Disponible en página web:

Todas las investigaciones temáticas.

Ficha 9: Análisis jurídicos



Descripción:

Se trata de análisis técnico-jurídicos especializados, sobre temas específicos, que estudian el derecho (como hecho, como norma y como valor). Los informes son documentos cortos que contienen una opinión legal sobre un tema. Estos análisis fueron colaboraciones gratuitas.

Documentos disponibles en página web:

Documentos citados en el cuadro 1A.5.

Recolección documental

La recopilación documental incluye el acopio de información proveniente de fuentes primarias (registros administrativos) y secundarias, sean textos o datos estadísticos. Este esfuerzo fue muy importante dentro del trabajo de campo, pues en el caso de los registros administrativos permitió acopiar información relacionada con 50 indicadores pertenecientes 28 de los estándares de la Auditoría. La recopilación bibliográfica allegó estudios relacionados con 64 indicadores correspondientes a 25 estándares.

Ficha 10: Registros administrativos



Descripción:

Los registros administrativos contienen la información, generalmente sistematizada y corroborada, que las instituciones públicas y privadas generan sobre sus actividades; por ejemplo, casos tramitados, eficiencia y resoluciones emitidas. Los registros administrativos sirvieron como referencia para los paneles de expertos y dieron fundamento empírico sistemático a los métodos interpretativos (cuadro 1A.6).

La compilación de los registros administrativos se efectuó mediante dos procedimientos

CUADRO 1A.5
Análisis jurídicos realizados

Autor	Tema
Fernando Castillo	Información sobre financiamiento público a los partidos políticos.
Hugo A. Muñoz	Protección de derechos: acceso a medios jurídicamente establecidos.
Laura Chinchilla	Protección de derechos: acceso igualitario a la justicia e independencia del Poder Judicial.
Fernando Guier	Medios de comunicación colectiva y normativa legal.
Sergio Vinocour	Delitos de cuello blanco y acceso a medios para combatir la corrupción.
Aldo Milano	Trasgresión a las normas que rigen la prestación de los servicios públicos.
Fernando Castillo	Mecanismos para la participación ciudadana en la formación de políticas públicas.
José Manuel Núñez	Legislación en materia de participación ciudadana.
Ronny Rodríguez	Financiamiento de partidos políticos.
Ana Virginia Calzada	Rendición de cuentas y recurso de amparo en políticas públicas.
Rocío Cerdas	Libertad de expresión de los diputados en el proceso legislativo
Francisco Dall'Anesse	Sujeción a la ley de representantes, funcionarios de confianza y servidores públicos.
Jaime Ordóñez	Reserva de ley y ámbito de los decretos ejecutivos.
Magda Inés Rojas	Las facultades legislativas del Poder Ejecutivo.
Jorge Calderón	Relaciones entre el Ejecutivo y Asamblea Legislativa.
Maruja Chacón	Campañas electorales y protección de derechos.
Hugo A. Muñoz	Políticas públicas.
Roxana Arroyo	Protección jurídica de los derechos de las mujeres.
Lena White	Protección jurídica a grupos étnicos y de género.
Rubén Chacón	La protección de los derechos de los indígenas.
Rodrigo A. Carazo	Protección jurídica a los derechos de las minorías étnicas.

CUADRO 1A.6
Principales registros administrativos empleados

Fuente	Tipo de registro o base de datos
Casa Presidencial	Archivo de decretos ejecutivos.
Asamblea Legislativa	Estadísticas legislativas.
Poder Judicial	Estadísticas judiciales.
Tribunal Supremo de Elecciones	Registro de denuncias por irregularidades electorales.
Defensoría de los Habitantes	Base de datos de quejas de los habitantes.
Sala Constitucional	Base de datos de votos por sentencia y derecho afectado.
Ministerio de Trabajo y Seguridad Social	Registro de organizaciones sociales.
Dirección Nacional de Desarrollo Comunal (DINADECO)	Registro de organizaciones comunales.
Registro Nacional	Registro de asociaciones civiles.
Dirección General de Servicio Civil	Registros de empleo público.
Partidos políticos	Registro de asambleas, normas y procedimientos.
Medios de comunicación colectiva	Bases de datos de noticias de <i>Al Día</i> y <i>La Nación</i> .

complementarios. En el primero, la compilación se hizo como parte de un estudio sobre un tema específico (v.gr. una ponencia especializada, una investigación temática); en el segundo, la compilación se realizó como un estudio independiente, orientado a identificar, verificar y valorar información de interés para los propósitos de la Auditoría.

Créditos:

Personal del equipo técnico del Proyecto Estado de la Nación (Elisa Sánchez, Alberto López, Rosslyn Jiménez, Vera Brenes y Evelyn Villarreal) tuvo a su cargo la recopilación de información.

Erick Jiménez y Angélica Chinchilla levantaron la información del Registro Nacional sobre asociaciones y Laura Arguedas la de sociedad civil.

Estos registros administrativos se encuentran disponibles en las oficinas del Proyecto Estado de la Nación, en la modalidad de hoja electrónica.

Documentos disponibles en página web:

Registro reprocesado de organizaciones de la sociedad civil.

Ficha 11: Recopilación bibliográfica y de estudios previos

01 OTRAS
INVESTIGACIONES

Descripción:

El trabajo de campo recopiló una bibliografía muy amplia de las publicaciones relacionadas con los temas de la Auditoría, tanto de autores costarricenses como de extranjeros; en total se reunió cerca de 1.500 fichas debidamente levantadas. La bibliografía se sistematizó mediante una base de datos especializada, Thesaurus, que incluye referencia a videocintas, casetes y sitios en internet.

Las fuentes de las que se extrajo la información son bibliotecas públicas ubicadas en:

- Las cuatro universidades públicas, incluyendo todos sus centros regionales.
- La Corte Suprema de Justicia.
- La Asamblea Legislativa.
- La Biblioteca Nacional.
- Archivos Nacionales.
- La Corte Interamericana de Derechos Humanos.
- El Centro de Documentación del Instituto Interamericano de Derechos Humanos.
- Bibliotecas en el exterior (por ejemplo, Universidad de Texas en Austin y la Biblioteca del Congreso de los Estados Unidos).

Créditos:

Lucrecia Molina, bibliotecaria del Instituto Interamericano de Derechos Humanos.

La base de datos se encuentra disponible en las oficinas del Proyecto Estado de la Nación.

Proceso de evaluación

Por proceso de evaluación se entiende el modelo, las normas, los criterios y los procedimientos que paneles de evaluadores especializados emplearon, en su condición de expertos, para evaluar el grado de cumplimiento de las aspiraciones democráticas -o estándares de alta calidad- en temas específicos de la vida política del país, con base en la información recopilada durante el trabajo de campo. Este proceso de evaluación tomó como referencia la propuesta original de ejecución de la Auditoría ciudadana (Gutiérrez y Vargas-Cullell, 1998).

El planteamiento original del modelo de evaluación puede ser consultado en la página web www.estadonacion.nu.or. En general, el proceso de evaluación respondió a este planteamiento pero, como ocurre con cualquier propuesta metodológica, en la práctica fue necesario efectuar ciertas modificaciones para ajustarlo a la disponibilidad real de tiempo, recursos financieros y humanos. En esta sección se sintetiza los principales contenidos del modelo de evaluación, se describe las modificaciones al planteamiento original y las razones que originaron tales cambios. Cabe advertir que el detalle de las valoraciones, críticas y sugerencias formuladas por cada uno de los 10 paneles de evaluación se presenta en los capítulos correspondientes, en la tercera parte de este informe.

Introducción

La evaluación de los hallazgos del trabajo de campo fue la última fase en la ejecución de la Auditoría ciudadana (véase el Prólogo de este informe). Se basó en un modelo con referencia a criterios, pues lo que interesó fue valorar el desempeño del país en áreas específicas de su vida política, de acuerdo con ciertos estándares de calidad democrática previamente definidos por el Consejo Consultivo de la Auditoría⁸.

El proceso comprendió dos actividades: por una parte, la desagregación analítica de los estándares (un movimiento de lo general a lo específico), con el fin de establecer criterios claros para la evaluación de los resultados; y, por otra, un proceso de agregación de las valoraciones de las y los expertos evaluadores para poder arribar a juicios globales sobre los estándares (es decir, un movimiento de lo específico a lo general). La desagregación analítica de los estándares de alta calidad democrática implicó la definición de variables, indicadores y condiciones de evaluación, y proporcionó

Proceso de agregación y desagregación de la Auditoría



RECUADRO 1A.1

Naturaleza de la evaluación de la Auditoría ciudadana

En su sentido más básico, la Auditoría es una deliberación (conversación, diálogo) entre ciudadanos sobre la democracia costarricense. Ciertamente, es una deliberación que se atiene a reglas formales para arribar a un fin específico: emitir una opinión informada sobre la calidad de la democracia, con repercusiones en el debate de los asuntos públicos del país. Se trata de un diálogo altamente formalizado. Mediante estas reglas formales se tiene acceso a una información común y se establece las condiciones y criterios para la evaluación, con el propósito de facilitar el diálogo y garantizar la máxima precisión posible en el significado de las calificaciones y de los acuerdos y desacuerdos entre los evaluadores. Esto hace distinta a la Auditoría de una conversación casual. Pero no por ello deja de ser una deliberación. Y este reconocimiento es muy importante para entender el alcance y limitación de los juicios evaluativos que realiza la Auditoría ciudadana. En primer lugar, la evaluación sobre la calidad democrática se ajusta a los indicadores, las preguntas de evaluación y los estándares definidos. No puede extrapolarse a otros temas que no fueron cubiertos por las investigaciones. En segundo lugar, las "notas" son juicios intersubjetivos, es decir, criterios elaborados en común por dos o más personas luego de revisar cierta

información, y no son calificaciones objetivas e irreversibles sobre la democracia costarricense. Finalmente, es evidente que paneles de investigación integrados de diferente manera podrían arribar a conclusiones distintas, incluso después de examinar la misma información. El cuidado por conformar paneles con un equilibrio de personas de diversos sectores sociales y políticos procuró minimizar los sesgos a favor de una u otra corriente de pensamiento, pero no contrarresta esa situación. Que paneles distintos puedan arribar a conclusiones distintas es un hecho cierto e inevitable, pero no sólo para una Auditoría ciudadana. Lo mismo ocurre con las percepciones y juicios de las personas, o con los estudios científicos sobre procesos sociales. No podría evitarse este riesgo aún cuando en una evaluación participaran todos los habitantes del país.

¿Qué aportan, entonces, los juicios evaluativos de los paneles? Mucho. Son juicios informados sobre temas relevantes para la democracia costarricense; indagan más allá de los prejuicios sobre esta democracia; constituyen una nueva forma de deliberación pública; tienden puentes entre ciudadanos de diversa procedencia social y política; estimulan el debate crítico y respetuoso y ayudan a desarrollar una nueva práctica de participación en los asuntos públicos y de petición de cuentas.

los criterios de evaluación a las y los expertos. Para la agregación de sus calificaciones, se diseñó una escala de evaluación ordinal de cinco grados para calificar el grado de cumplimiento percibido de un estándar y un método que permite sumar resultados, desde los indicadores hasta los estándares. Se formó 10 paneles de evaluación, cada uno integrado por al menos 3 personas y a cargo de un tema específico⁹.

En el diseño original de la fase de evaluación de la Auditoría se planteó la participación de tres instancias en el proceso: el equipo técnico, los paneles de evaluación y una junta consultiva cuya función sería resolver consultas de los evaluadores, sustituir panelistas en caso de renuncia y, en general, tomar decisiones sobre la marcha de la evaluación. En la práctica, esta junta consultiva no operó y el equipo técnico asumió esas funciones¹⁰.

Modelo de evaluación

El modelo de evaluación de la Auditoría contiene la especificación de un conjunto de reglas y criterios comunes para todos los paneles evaluadores. Su propósito fue asegurar que las calificaciones individuales y colectivas de los expertos sobre la calidad democrática se establecieran como resultado de procesos similares y mediante la aplicación de escalas comunes de evaluación. Estas calificaciones son valoraciones sobre cuán cercana o lejana se encuentra la vida democrática en Costa Rica en relación con un horizonte de buenas prácticas (USCD). Como se señaló en el Prólogo, el USCD, formado por 33 estándares o aspiraciones

de alta calidad democrática y su desagregación en 138 variables, fue aprobado por el Consejo Consultivo de la Auditoría en agosto de 1998, luego de un proceso de consulta y discusión¹¹.

Es importante reiterar que el punto de partida de la evaluación fueron los estándares o aspiraciones democráticas. Los estándares son los criterios para evaluar la vida democrática. De acuerdo con la International Organization for Standardization (ISO) un estándar es “un acuerdo documentado que contiene especificaciones técnicas u otros criterios específicos para ser empleados de manera consistente como reglas, guías o definiciones de características” que permiten juzgar si una cosa es apropiada para el fin que se persigue (www.iso.com). En síntesis, un estándar es un acuerdo documentado sobre una buena práctica democrática que, de existir en Costa Rica, permitiría afirmar que en el país existe una alta calidad democrática en esa materia.

En la evaluación, así como en la investigación, la identificación de las variables fue el primer paso para la evaluación de los estándares. Una variable es un componente o atributo de una aspiración democrática, definido como tal por el Consejo Consultivo. En la Auditoría existen variables unidimensionales y multidimensionales. Una variable es unidimensional cuando mide una sola característica, y multidimensional cuando mide dos o más características. Cada dimensión de una variable fue medida por un indicador distinto. Así, las 138 variables fueron desagregadas en 240 indicadores.

Para efectos de la evaluación de la calidad democrática, la desagregación analítica de cada uno

CUADRO 1A.7

Ejemplos de un indicador de comprobación y un indicador de evaluación

Tipo de indicador	Nombre del indicador	Definición	Tipo de valoración
Comprobación	Existencia de mecanismos de consulta pública en la aprobación del gasto público	Los procedimientos establecidos para la aprobación del gasto público por las instancias respectivas incluyen un plazo definido para la convocatoria y realización de audiencias públicas con expertos, interesados y público en general, para manifestarse sobre las iniciativas.	No se cumple: revisión de la legislación y el Reglamento de la Asamblea Legislativa
Evaluación	Ejercicio de censura por parte de medios de comunicación colectiva	Los medios de comunicación colectiva se abstienen de ejercer censura previa sobre las manifestaciones de las personas que recurren a ellos o son entrevistados por ellos.	Mayoría del panel: cumplimiento general Minoría del panel: cumplimiento medio Con base en el trabajo de campo.

FUENTE: PROYECTO ESTADO DE LA NACIÓN, 2000.

Indicadores de comprobación, de evaluación e indicadores descartados



FUENTE: PROYECTO ESTADO DE LA NACIÓN, 2000.

de los niveles de análisis antes citados (estándares, variables e indicadores) implicó tres pasos: agrupación de estándares en dominios temáticos, transformación de las variables en preguntas de evaluación y, finalmente, la definición de los indicadores y su especificación en condiciones de evaluación.

Las preguntas de evaluación

En la evaluación, las variables fueron convertidas en preguntas. Estas preguntas son las interrogantes que los expertos debían calificar, una especie de guión de evaluación. La mayoría de los paneles debió contestar entre 10 y 16 preguntas, excepto en los temas de la calidad democrática de la administración de la justicia, de la opinión pública y de los gobiernos locales. En los dos primeros casos, cada panel debió responder una cantidad superior de preguntas; en el último, sólo 4.

La redacción de las preguntas fue, por lo general, fiel al texto aprobado por el Consejo Consultivo como desagregación del estándar. Sin embargo, hubo tres tipos de modificaciones que conviene señalar:

- Modificaciones de estilo para contar con preguntas claramente redactadas. En estos casos, no se alteró el sentido original.
- Modificaciones de fondo, cuando una revisión cuidadosa detectó imprecisiones o errores evidentes en el texto original, que no habían sido advertidos. En estos casos, el equipo técnico eliminó el error, o enmendó la imprecisión tomando como referencia el texto del estándar y procurando una reformulación apegada al sentido del estándar. Cabe señalar que 10 de las 138 variables estaban en esa situación.

- Eliminación de la variable cuando, al ser transformada en pregunta de evaluación, el equipo técnico determinó que era redundante; 3 variables de las 138 tenían esa condición.

Especificación de los indicadores de evaluación

Cada uno de los 240 indicadores fue enunciado y definido. El objetivo fue proporcionar a los evaluadores códigos de entendimiento comunes y

RECUADRO 1A.2

Ejemplo de las piezas de información disponibles para un indicador

El indicador de “elecciones municipales limpias” tenía 4 piezas de información provenientes de 4 métodos distintos:

- Número de denuncias por fraude electoral o irregularidades presentadas ante el Tribunal Supremo de Elecciones (registros administrativos).
- Número de quejas planteadas ante los medios de comunicación colectiva por fraude electoral o irregularidades (registros administrativos).
- Porcentaje de personas que creen que las elecciones en Costa Rica son limpias (encuesta nacional).
- Percepciones de grupos focales sobre la limpieza de los comicios (grupos focales).

En estos casos, el equipo técnico identificó, ordenó y aseguró que las y los evaluadores recibiesen todas estas piezas de información como parte del material de evaluación.

precisos para la revisión del material empírico y su evaluación posterior. Esta minuciosa labor requirió rondas sucesivas de discusión crítica dentro del equipo técnico, con el fin de asegurar que la definición de cada indicador fuese lo más clara posible, consistente con la pregunta de evaluación y coherente con las definiciones de otros indicadores similares.

Los indicadores fueron clasificados en dos grupos. El primero fue denominado “indicadores de comprobación”, pues incluye los indicadores que refieren a la existencia o no de una característica deseable y no al grado (o manera) en que ésta existe. En estos casos no es necesaria la evaluación de un panel de expertos, pues basta la comprobación de su existencia (o inexistencia) para examinar cómo se comporta el indicador. El desempeño del país en los temas referidos por indicadores de comprobación fue verificado por el equipo técnico de la Auditoría.

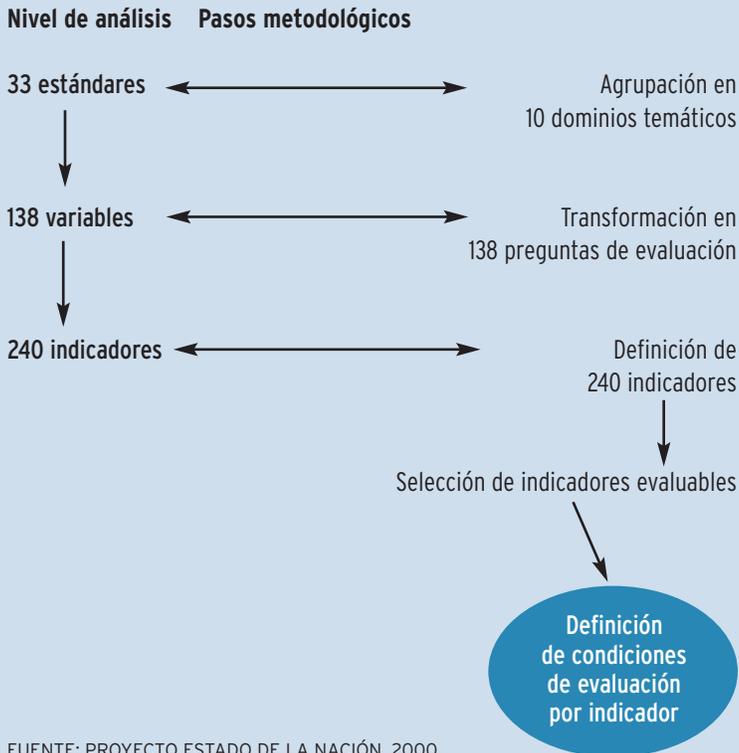
El segundo grupo fue denominado “indicadores de evaluación”, pues incluye todos los indicadores que refieren al grado (o manera) en que una característica deseable existe en Costa Rica. Estos

fueron los indicadores sometidos a la valoración de los paneles de expertos.

El siguiente paso fue examinar la información disponible para cada indicador. Esta revisión permitió determinar que, para ciertos indicadores, no existía información o que ésta no se relacionaba directamente con el tema. En estos casos, el equipo técnico realizó un último esfuerzo de búsqueda e identificación de información para subsanar vacíos. Concluido este trabajo, se cotejó de nuevo la información con los indicadores y se descartó aquellos indicadores sin información (cuadro 1A.7).

CUADRO 1A.9

Representación sintética del proceso de desagregación analítica para la etapa de evaluación



FUENTE: PROYECTO ESTADO DE LA NACIÓN, 2000.

RECUADRO 1A.3

Problemas prácticos con los indicadores y condiciones de evaluación

A lo largo del proceso de evaluación de la Auditoría, los paneles detectaron problemas con ciertos indicadores y condiciones de evaluación. La especificación de 210 indicadores y de más de 450 condiciones de evaluación se realizó en un período muy corto, debido a los apremiantes plazos del cronograma de trabajo. Ello impidió que se realizara un debido proceso de verificación de la consistencia y confiabilidad de los indicadores y las condiciones con grupos de expertos. En dos paneles, el de cultura cívica y el de opinión pública, las y los evaluadores fueron especialmente críticos a este respecto. Por otra parte, en todos los paneles hubo debates, más o menos prolongados, sobre la jerarquía de las fuentes de información.

En los capítulos del informe se consigna los casos en que los evaluadores formularon críticas metodológicas. El hecho de que el modelo de evaluación fuese definido cuando ya había avanzado el proceso de ejecución de la Auditoría contribuyó al surgimiento de problemas de consistencia de los indicadores y las condiciones de evaluación. Es necesario reconocer que, por ser esta una primera experiencia de exploración de la calidad democrática, al inicio no se tenía una idea precisa sobre la complejidad metodológica y técnica de una evaluación participativa. La principal lección aprendida es que, en una próxima experiencia, la especificación del modelo de evaluación debe realizarse desde el principio, junto con el diseño del trabajo de campo.

Finalmente, la decisión de no jerarquizar *a priori* las fuentes de información y dejar esta discusión como una tarea propia del panel ciertamente generó debates sobre su pertinencia, pero, en general, facilitó la discusión de los temas sustantivos entre las y los evaluadores. De esta manera ellos tuvieron que crear criterios de interpretación comunes (cuando fue posible), en vez de limitarse a aplicar criterios previamente establecidos.

Escala de evaluación aplicada por los paneles de evaluadores

Escala	Categoría	Definición
5	Cumplimiento pleno	En el indicador bajo examen se cumple plenamente la aspiración de alta calidad democrática y no se detecta problemas.
4	Cumplimiento general	En el indicador bajo examen se cumplen en general la aspiración de alta calidad democrática; se detecta problemas, pero éstos tienen una importancia secundaria.
3	Cumplimiento medio	En el indicador bajo examen existe un balance de logros e incumplimientos que crean una equidistancia entre el cumplimiento pleno y un incumplimiento severo.
2	Incumplimiento general	En el indicador bajo examen se incumple en general la aspiración de alta calidad democrática; se detecta logros, pero éstos tienen una importancia secundaria.
1	Incumplimiento severo o frontera crítica	En el indicador bajo examen se incumple severamente la aspiración de alta calidad democrática y no se detecta logros.

Nota: el nombre de las categorías es ligeramente distinto al consignado en el modelo original de evaluación. Sin embargo, no cambió el tipo de escala ni la definición de cada categoría.

FUENTE: ADAPTADO DE PROYECTO ESTADO DE LA NACIÓN, 2000.

Medición del desempeño de los indicadores: las condiciones de evaluación

La revisión de la información disponible para cada indicador arrojó que, en muchos casos, existía material proveniente de distintos métodos de investigación. Cada fuente ofrecía datos distintos, que no podían ser sumados u ordenados por medio de una jerarquía *a priori* sin recurrir a supuestos difíciles de defender (recuadro 1A.2). En ningún caso se supuso que una percepción era más (o menos) importante que una denuncia ante el Tribunal Supremo de Elecciones, o que una queja ante un medio de comunicación, pues se carecía de un fundamento teórico para hacerlo. La jerarquización de la información fue, precisamente, parte de las discusiones de los paneles de expertos a la hora de formular sus calificaciones.

“No hay manera de cruzar o combinar los criterios subjetivos y objetivos de forma que sean mutuamente armónicos e intersubjetivamente unívocos. Imponer sólo criterios objetivos sería claramente paternalista. Por el otro lado, aceptar sólo criterios subjetivos es prohibirse examinar y criticar cómo las preferencias reveladas (de las personas) pueden ser causadas por un bajo

desarrollo humano y por una ciudadanía de baja intensidad. No podemos, entonces, dejar de adoptar ambos criterios” (O’Donnell, 2000).

El procedimiento fue, primero, identificar las fuentes de información y, luego, suponer que la información que cada una arrojaba era independiente. Independiente significa que cada pieza o “celda” de información observada por alguno de los métodos de investigación era valorada por aparte y en relación consigo misma. Esta, la unidad máxima de desagregación, fue denominada “condición de evaluación”. Cada indicador fue medido por dos o más condiciones de evaluación (el rango osciló entre 2 y 5). En total se especificó más de 450 condiciones de evaluación.

Con la definición de las condiciones de evaluación se concluyó el proceso de desagregación analítica de los estándares. Las y los evaluadores tuvieron así un “puente” que conectaba cada parte de información con un indicador, pregunta de evaluación y aspiración o estándar democrático. Esta desagregación es, por así decirlo, el mapa de los vasos comunicantes entre los procesos de investigación y evaluación de la Auditoría ciudadana (cuadro 1A.8).

Razones por las que se varió la función del juicio con base en criterios definidos durante la fase final de la evaluación

En el modelo original de evaluación se estipulaba que, durante la fase colectiva de trabajo, las y los evaluadores revisaban el juicio compartido “para asegurar la correcta aplicación de las condiciones de evaluación” antes de iniciar la discusión tendiente, de ser posible, a arribar a un juicio de experto consensual. En la práctica se aplicó un procedimiento distinto. En la fase colectiva no se revisó la calificación de los juicios compartidos, sino que, con base en los resultados de las calificaciones individuales de experto, la discusión se desarrolló sobre los indicadores en los que no existía acuerdo, con el fin de arribar a un juicio colectivo de experto.

Las razones que justificaron esta modificación fueron las siguientes:

- La revisión de los juicios compartidos era un proceso lento, engorroso e implicaba una inversión de tiempo considerable de las y los evaluadores. Era reabrir el proceso desde el inicio.
- La parte pendiente era la revisión de los desacuerdos, para determinar si éstos obedecían a diferencias profundas e insalvables de criterio, o a matices subsanales. Era allí donde debía concentrarse la discusión.
- La función principal del juicio compartido había sido establecer un terreno común de evaluación, cosa que sí hizo, y servir de guía para la lectura de los hallazgos del trabajo de campo. Aplicar el juicio compartido en etapas posteriores era encasillar a las y los evaluadores y reducirlos a verificadores del cumplimiento de las condiciones de evaluación definidas por el equipo técnico. Ello significaba desaprovechar su conocimiento experto.
- La opción escogida fue pedir a las y los evaluadores comentarios críticos sobre el juicio compartido, o aclaraciones a comentarios escritos por ellos en los cuadernos de evaluación.

La calificación de la calidad democrática

Escala de evaluación

Las y los evaluadores aplicaron una escala ordinal de 5 grados en los juicios que emitieron. En todos los casos, la escala respondió a la siguiente pregunta: ¿se cumple con el estándar de alta calidad democrática en el indicador bajo examen? Las 5 categorías

definidas para responder esta interrogante corresponden a una escala de Lickert (cuadro 1A.9).

Tipos de valoración

El proceso de valoración de los hallazgos del trabajo, con base en la escala mencionada, comprendió dos tipos de juicio por parte de los evaluadores. El primero fue el juicio con base en **criterios predefinidos** (o juicio compartido). Se realizó a partir de una especificación de las condiciones de evaluación formuladas por el equipo técnico, comunes a todos los evaluadores y expresamente aceptadas por ellos. Las personas asignaron una calificación ciñéndose a los criterios y normas entregadas. En este caso, su labor evaluadora se caracterizó por la lectura y verificación de la información recibida y su ubicación dentro de la escala que le fue entregada. El punto de partida fue la valoración de cada una de las condiciones de evaluación establecidas para un indicador y, luego, la aplicación de un método simple de agregación de puntaje para establecer una calificación por indicador¹². El segundo tipo de juicio fue el **juicio de experto** y se realizó con base en los conocimientos y percepciones de cada evaluador. Aquí la persona tuvo libertad para especificar sus criterios y asignar una calificación que podía o no ser concordante con la del juicio compartido. En síntesis, en la valoración los expertos aplicaron el juicio compartido y el de experto.

En el planteamiento original del modelo de evaluación se indicaba que el juicio con base en criterios definidos, o juicio compartido, era “la base principal de la evaluación de la vida democrática” y que el juicio de expertos tenía “una función complementaria al juicio compartido; puede modificarlo pero no anularlo completamente” (Proyecto Estado de la Nación, 2000). Para expresar la calificación total otorgada por un evaluador en un indicador, se aplicó una fórmula que dio más peso al juicio compartido que al juicio de experto, de manera que para modificar sustancialmente la calificación del juicio compartido, el evaluador debía expresar una importante discrepancia en su juicio de experto¹³. Esta fórmula de agregación fue conocida por los evaluadores en la sesión de inducción, pero le correspondió al equipo técnico aplicarla:

$$CF = JC + (JC-JE) / 3$$

donde JC = calificación obtenida en el juicio compartido; JE = calificación obtenida en el juicio de experto y CF = calificación final de evaluador. El resultado se redondea al número entero más cercano.

El proceso seguido fue un tanto distinto. Este método de agregación se aplicó a todas las calificaciones individuales de los evaluadores; ello permitió

Procedimiento analítico y metodológico de la evaluación

Agregación de resultados por el equipo técnico

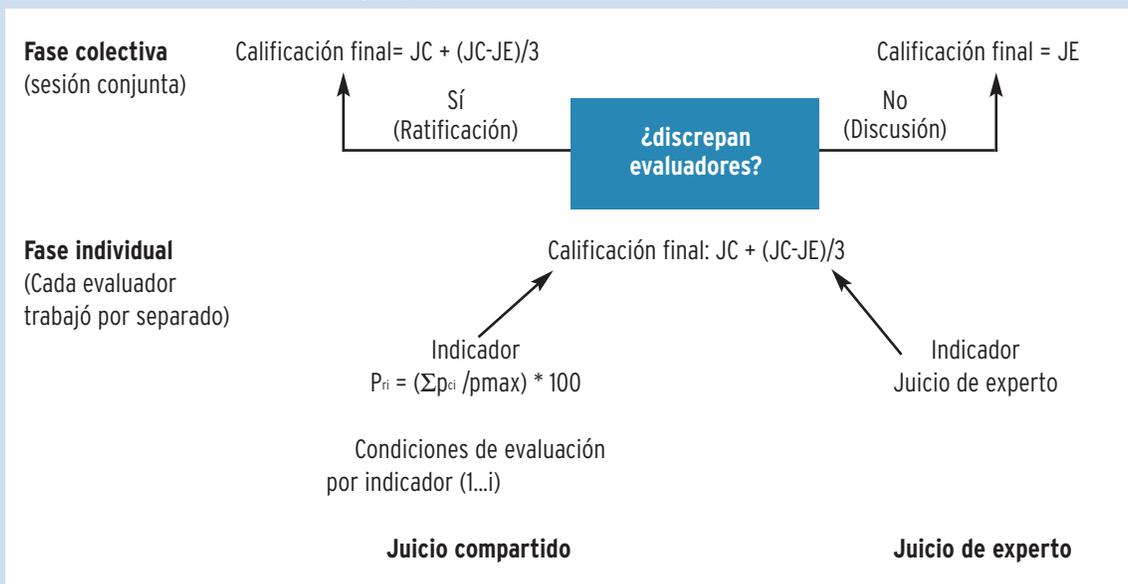
Aspiración democrática

$$Prad = \sum Pre/E$$

Pregunta de evaluación

$$Pre = \sum Pri/I$$

Evaluadores y equipo técnico: evaluación por indicador (Pri)



FUENTE: PROYECTO ESTADO DE LA NACIÓN, 2000.

compararlas entre sí, pues todas las personas habían aplicado los mismos criterios y escalas de evaluación. Sin embargo, contra lo previsto, en la fase de trabajo colectivo del panel, el único juicio aplicado fue el juicio de experto (recuadro 1A.4). La razón principal fue una reevaluación del alcance y la pertinencia del juicio compartido en la última fase de la evaluación.

Los desacuerdos entre panelistas

Los acuerdos y desacuerdos entre las y los panelistas sobre la calificación de un indicador eran resultados igualmente importantes para la Auditoría. Por ello, a las y los panelistas se les reiteró la solicitud de que no forzaran “consensos” o negociaran calificaciones cruzadas (aceptar una calificación desfavorable en un asunto, para lograr de los demás una calificación favorable en otro).

En el informe final se consigna escrupulosamente los casos en que la calificación grupal fue

producto de calificaciones discrepantes. Además, se insertó recuadros de texto en los cuales se resumen los motivos y razonamientos relacionados con esas discrepancias.

Agregación de calificaciones

La evaluación procedió en un sentido contrario al proceso de desagregación analítica descrito en una sección anterior. En efecto, dentro de cada dominio temático, los evaluadores valoraron cada uno de los indicadores. Luego, mediante un sencillo procedimiento aritmético, el equipo técnico agregó los resultados en cada pregunta de evaluación y, finalmente, en cada aspiración democrática. La base del proceso de evaluación fue, entonces, la valoración de los indicadores y no el estándar democrático. Se trató de una calificación modular: una vez valorados los indicadores, las personas no tuvieron que realizar nuevas rondas de calificación de las preguntas de evaluación y los

estándares democráticos. El equipo técnico sumó los resultados obtenidos en el nivel más bajo de la escalera de abstracción. El procedimiento analítico seguido se presenta en forma resumida en el cuadro 1A.10. En los casos en que había desacuerdos entre los panelistas sobre la calificación de un indicador, este método de agregación permitió arribar a calificaciones grupales sobre las preguntas de evaluación y los estándares.

Fases de trabajo del proceso de evaluación

Las fases de trabajo del proceso de evaluación fueron las siguientes:

- Preparación del material de evaluación. Esta etapa comprendió la clasificación de las investigaciones

del trabajo de campo por aspiración y variable, y la elaboración de un cuaderno de evaluación con las instrucciones, indicadores y condiciones de evaluación para las y los evaluadores. De esta manera, cada experto recibió un archivo con la información de base y un cuaderno de evaluación para calificar el desempeño de los indicadores. En esta etapa, como ha sido señalado, el principal problema enfrentado fue la insuficiente verificación de la consistencia y confiabilidad de ciertos indicadores de evaluación.

- Conformación de los paneles de evaluadores. Se formaron 10 paneles de evaluadores, en los que participaron 34 personas (véase Reconocimientos). Esta etapa implicó un arduo trabajo, pues el

CUADRO 1A.12

Síntesis del proceso de evaluación

Etapa	Cambio respecto a diseño original	Comentario	Duración original	Duración real
Preparación del material de evaluación	Ninguno	La clasificación del material hizo que su lectura fuera difícil. Poco tiempo para verificar consistencia y confiabilidad de condiciones de evaluación	No prevista	3 meses
Conformación de instancias	No se formó la junta consultiva		No prevista	
Sesión de inducción	Ninguno	Fue necesario combinar sesiones colectivas con sesiones individuales		
Conformación definitiva de paneles	No se cumplió en el plazo previsto	Lento y difícil encontrar sustitutos	1 semana	3 semanas
Trabajo individual de evaluadores	No se cumplió con el plazo previsto	Subestimación de la complejidad de la tarea	1 semana	4 semanas (promedio)
Procesamiento de resultados individuales	No se cumplió con el plazo previsto	Consultas con evaluadores demoraron el procesamiento	1 semana	2 semanas
Sesiones colectivas	No se aplicó juicio compartido. Debate sobre desacuerdos	Fue necesario aprovechar más el juicio de experto y optimizar el poco tiempo disponible	No prevista	5 semanas
Elaboración de dictámenes	No se cumplió con el plazo previsto	Se substituyó la elaboración de dictámenes por redacción de capítulos	1 semana	6 meses
Retroalimentación del proceso	Se realizó como parte de las sesiones colectivas		No prevista	5 semanas

equipo técnico debió persuadir a personalidades políticas y académicas para que aceptaran participar *ad honorem* en un ejercicio que implicaba la inversión de una buena cantidad de tiempo (en promedio, unos dos días completos de trabajo, además de la inducción y la sesión de trabajo colectivo).

- Sesión de inducción de las y los evaluadores. En esta etapa se convocó y se instruyó a las y los evaluadores sobre las normas y procedimientos de evaluación, se realizaron ejercicios prácticos y se fijó un cronograma de trabajo. La única variante con respecto al planteamiento original consistió en que con algunos evaluadores la actividad fue individual, pues no pudieron asistir a las dos sesiones de inducción convocadas por el equipo técnico.
- Conformación definitiva de los paneles de evaluación. La diferencia con respecto al planteamiento original fue que, al no crearse una junta consultiva, los cambios fueron decididos por el equipo técnico. Como estaba previsto, fue necesario realizar cambios de última hora en la integración de los paneles. La sustitución de panelistas fue un proceso lento y dificultoso. En un caso, no hubo otro remedio que permitir una excepción a la regla de que las personas que hubiesen realizado una contribución puntual no evaluaran el tema en el cual su contribución formaba parte del material de evaluación¹⁴.
- Fase individual de trabajo de los evaluadores. En esta etapa, las y los evaluadores revisaron el material de evaluación, procedieron a calificar el desempeño de los indicadores (juicio compartido y juicio de experto) y formularon sus comentarios. Concluyó con el envío de los cuadernos de evaluación, debidamente llenos, de vuelta al equipo técnico. Originalmente se pensó que esta etapa tendría una duración de aproximadamente 7 días naturales, pero fue una estimación equivocada, pues las y los evaluadores tardaron en promedio 4 semanas en realizar esta tarea y hubo casos en que el proceso tomó 6 semanas. La principal razón fue una seria subestimación de la complejidad de la labor de los evaluadores (cuadro 1A.11).
- Procesamiento de los resultados de la calificación individual de las y los evaluadores. El equipo técnico revisó cuidadosamente los cuadernos de evaluación recibidos, consultó con los evaluadores cuando existían dudas sobre los comentarios

o la calificación de un indicador, creó una hoja electrónica para procesar los resultados, “vacío” en estas hojas los resultados y creó un cuadro-resumen para orientar el trabajo de las sesiones colectivas de los paneles.

- Fase de trabajo colectivo de los paneles de evaluación. Las sesiones de trabajo se realizaron de conformidad con el diseño del proceso de evaluación, con la excepción ya comentada: en esta fase no se aplicó el juicio compartido. En todos los paneles, las sesiones colectivas fueron actividades que dieron origen a discusiones en profundidad sobre la calidad de la democracia en Costa Rica y a observaciones sobre el proceso de evaluación seguido. A juicio del equipo técnico, esta fue una actividad clave de la Auditoría.
- Elaboración de los capítulos. En el diseño original del proceso de evaluación se tenía prevista la elaboración de dictámenes cortos, en los que se sintetizara los principales hallazgos de la investigación de cada estándar, las calificaciones y comentarios de los evaluadores. En la práctica, esta etapa fue muy distinta: en vez de dictámenes, el equipo técnico elaboró capítulos más extensos, cada uno de los cuales presenta con cierto detalle el material de evaluación de cada panel y los resultados del proceso. Varias razones explican este cambio en el procedimiento. La primera es que es que el procesamiento de las discusiones de los paneles fue una actividad delicada, pues debió desarrollarse un estilo de redacción que expresara certera y equilibradamente los acuerdos y desacuerdos entre los panelistas. La segunda razón es que la elaboración de las herramientas para presentar sintéticamente los hallazgos y calificaciones de los paneles consumió más tiempo de lo previsto. Implicó varias aproximaciones sucesivas hasta obtener una versión satisfactoria. Finalmente, en el diseño original se subestimó severamente la complejidad de la tarea de sintetizar los principales hallazgos del trabajo de campo. En la práctica, el equipo técnico tuvo que revisar todo el material de evaluación, seleccionar los hallazgos de conformidad con las indicaciones de los panelistas y cuidar que la evidencia respaldara los resultados de las evaluaciones.

Preparación del informe final

La preparación del informe final de la Auditoría ciudadana se realizó entre enero y abril de 2001. El equipo técnico elaboró la versión preliminar de cada capítulo. Estas versiones preliminares

fueron colocadas en la página web www.estadonacion.or.cr. Se desarrolló una consulta privada con los 34 panelistas y 42 miembros del Consejo Consultivo, a quienes se les suministró una clave de acceso para que, con plazos de aproximadamente 21 días, comentaran estas versiones preliminares. Aunque todos podían comentar sobre cualquier capítulo, a las y los evaluadores se les pidió específicamente que revisaran el capítulo correspondiente a su panel.

El informe adoptó un estilo narrativo particular. Los 10 capítulos que forman la tercera parte del presente informe -su médula- tienen una organización modular:

- Una introducción que justifica la importancia del tema de la evaluación, un resumen de la valoración general del panel, los métodos de investigación empleados, las limitaciones y antecedentes relevantes.
- Los componentes principales de cada capítulo son los estándares de calidad democrática. A su vez, los estándares se dividen en las preguntas de evaluación.
- En cada estándar la presentación de resultados se organiza en dos secciones: a) Síntesis de resultados y evaluación del panel; b) Información de base.
- En la sección “Resumen de hallazgos y resultados de evaluación” se presenta cuatro cuadros y un gráfico: a) Resumen de las principales evidencias; b) Otros hallazgos relevantes; c) Cuadro

resumen con la valoración del panel sobre el estándar; d) Gráfico con las calificaciones del panel por indicador; e) Indicadores de comprobación (cuando existen).

- En la sección “Información de base” se presenta, con detalle, una selección de los resultados de investigación más relevantes para el tema en evaluación. Se divide por preguntas de evaluación, a cada una de las cuales corresponde un acápite separado. Los acápites están organizados a la manera de hipertextos. Se presentan “cápsulas de información”, cada una de las cuales sintetiza un hallazgo del trabajo de campo, independiente de cualquier otro y obtenido por medio de un método de investigación particular. Cada cápsula de información está precedida por un ícono que representa al método de investigación que la respalda.

En síntesis, el informe de la Auditoría ciudadana no es un texto convencional, en el cual el autor establece las conexiones lógicas entre párrafos y formula conclusiones. Es un “puntillero” de las piezas de información más relevantes que empleó el panel de evaluación respectivo para formular sus calificaciones y observaciones. Se trata de un texto que bombardea al lector pero, al mismo tiempo, le deja amplio margen para sacar sus propias conclusiones. Este estilo procura reflejar la condición esencial de la Auditoría: es un proceso y un producto de una conversación colectiva (no una interpretación acabada de autor) sobre la situación actual de la democracia costarricense.

Notas

1 Son tres las razones por las cuales el concepto de auditoría resultó útil a esta iniciativa:

- a) Así como las auditorías tradicionales hacen referencia a la rendición de cuentas (Sheldon, 1996; General Accounting Office, 1988), la evaluación de la vida democrática en Costa Rica procuró extender el alcance de la rendición de cuentas horizontal (O'Donnell, 1997).
- b) Así como las auditorías tradicionales establecen estándares desde los cuales se pueden juzgar las prácticas reales en una empresa o institución, el estudio sobre la calidad de la democracia costarricense aplicó estándares de evaluación.
- c) "Cualquier auditoría consiste de una serie de indagaciones, pruebas, y reportes. Las indagaciones se hacen a través de entrevistas, observaciones y estudios sobre la evidencia sustancial, para que el cliente logre tener una adecuada comprensión sobre la situación" (Sheldon, 1996). Igualmente, en la Auditoría ciudadana hubo un "cliente involucrado", las y los ciudadanos, e indagaciones basadas en una mezcla ecléctica de métodos de investigación y pruebas, que no conforman un ejercicio delimitado a ciertos temas.

2 No existe una relación de necesidad entre un método y otro. Esto explica que la unidad básica para el trabajo de campo no sea el dominio en el que se ubican los diferentes estándares, ni siquiera los propios estándares. La unidad base o guía de investigación fue cada uno de los indicadores de la Auditoría.

3 El diseño de los estándares (y de los propios indicadores) como módulos temáticos independientes tiene la ventaja de que la inconsistencia en el abordaje de un método no tiene un efecto negativo sobre el conjunto. Así por ejemplo, la falla de un investigador o colaborador en la realización de un análisis jurídico podría tener un peso específico muy grande sobre la calidad del estudio del indicador en el que se ubique, pero el problema no afecta o retrasa el tratamiento de otros estándares, o incluso el de otros indicadores del mismo estándar donde ocurre el problema.

4 Algunos indicadores, por su naturaleza, debían ser estudiados por más de una investigación temática o por más de una ponencia especializada. Es el caso, por ejemplo, de los indicadores relacionados con la participación ciudadana en las políticas públicas. Se recurrió a estudios jurídicos y administrativos sobre las oportunidades de participación que las instituciones públicas ofrecen a las y los ciudadanos, a la encuesta de opinión para conocer percepciones y actitudes de participación, a los grupos focales y a los registros administrativos. Hubo dos estándares completos -esto es, que incluyen todos sus indicadores- en los que se consideró importante que fuesen estudiados por la mayor cantidad de investigaciones: el estándar referido a la fuerte creencia de la ciudadanía en la democracia como valor en sí mismo y como un medio político adecuado para mejorar el bienestar público, y el estándar relacionado con la identificación de la ciudadanía como parte de una unidad cívica nacional.

5 La combinación de métodos intentó reducir las diferencias, inevitables, en la "profundidad" de la investigación y evaluación de los diversos estándares. En la propuesta de ejecución de la Auditoría, así como en las reuniones del panel académico asesor, fue expresamente señalado que no todos los estándares serían investigados con igual meticulosidad, pero que para todos los casos se acopiaría información novedosa y relevante.

6 En dos casos, los informes finales de investigación no estaban listos cuando se inició el proceso de evaluación. Las razones para el atraso fueron las siguientes:

- En el caso del ejercicio de los medios de comunicación, el diseño del ejercicio, la negociación con las empresas y la duración misma de la observación fueron un proceso bastante más dilatado que el previsto originalmente. Al momento de la evaluación, las y los evaluadores disponían de un avance esquemático de investigación (capítulo 12). Cabe señalar que los resultados finales de este informe fueron conocidos por el panel durante el proceso de redacción del capítulo respectivo y no modificaron las calificaciones ni los términos de debate.
- En el caso de los Sitios Centinela, al momento de la evaluación las y los evaluadores conocieron los resultados de una primera ronda de ejercicios que incluyó a 4 municipios. La segunda ronda se realizó en forma paralela al proceso de evaluación. La razón para este atraso fue que los recursos comprometidos para esta iniciativa fueron transferidos al Proyecto Estado de la Nación seis meses después de lo previsto. Al igual que en el caso anterior, los resultados finales de este informe fueron conocidos por el panel durante el proceso de redacción del capítulo respectivo y no modificaron las calificaciones ni los términos de debate.

7 En cada sesión participaron 6 ó 7 personas, según la lista de invitados suministrada por los responsables del proyecto. Esta forma de citar a las y los participantes es parte de la obligación que tiene la empresa Unimer, de no asociar resultados con la identidad de los participantes. Además, en este estudio 5 personas solicitaron expresamente que se tuviera un especial cuidado en garantizar la confidencialidad de sus opiniones, ya que con solo citar su cargo se les identificaría.

8 La decisión de adoptar un modelo con referencia a criterios fue recomendada por el panel de expertos en evaluación que acompañó el diseño y planeamiento de esta fase de la Auditoría. La otra opción era adoptar un modelo de evaluación con referencia a normas, mediante el cual se valora la posición relativa que ocupa una cosa observada dentro de un grupo (Montero, 1999).

9 En este punto afloran diferencias entre las auditorías a empresas o instituciones y la denominada "Auditoría ciudadana". La primera diferencia estriba en el hecho de que las empresas contratan firmas de auditoría que son externas a ellas. La Auditoría ciudadana, en cambio, no sólo procuró ser una herramienta en manos de los ciudadanos, sino también que ellos mismos participaran en su realización. La segunda diferencia reside en que la Auditoría ciudadana, además de fortalecer la rendición de cuentas dentro del sistema político, persigue también otro objetivo: contribuir al desarrollo de nuevas prácticas políticas entre los ciudadanos y las instituciones, con el fin de estimular la vida democrática. Pese a que el concepto de auditoría no incluye este aspecto, no se encontró uno mejor para expresar la naturaleza de la evaluación sobre la calidad de la democracia.

10 La principal razón por la cual la junta consultiva no pudo formarse fue que la conformación de los paneles de evaluación, las sesiones de inducción y el seguimiento a su trabajo resultó ser un proceso largo y más complejo que lo previsto. La organización de la junta era una tarea adicional, que desbordaba las capacidades del equipo técnico. Si bien esto no descarta la importancia de una junta consultiva, en la práctica su ausencia no fue un obstáculo para realizar la evaluación, al menos en el caso de Costa Rica.

11 La consulta y discusión en el Consejo Consultivo incluyó los siguientes pasos y se desarrolló entre junio y setiembre de 1998: a) conversación con cada una de las personas invitadas a formar parte del Consejo Consultivo, para explicar la naturaleza de la Auditoría y los alcances de la participación solicitada; b) entrega y llenado de una encuesta sobre posibles estándares democráticos por parte de las personas que aceptaron la invitación; c) procesamiento de la información; d) realización paralela de seis sesiones con grupos focales de ciudadanos "comunes y corrientes", lo que posteriormente permitió examinar la consistencia de las opiniones del Consejo Consultivo. En general, las opiniones fueron bastante coincidentes; e) primera especificación de estándares democráticos y su desagregación analítica (en este paso se contó con la asesoría del Prof. Michael Coppedge, de la Universidad de Notre Dame, Indiana); f) sesión plenaria del Consejo Consultivo para adoptar una decisión sobre los estándares, con la participación del Prof. Guillermo O'Donnell de la misma Universidad, quien acompañó y asesoró la Auditoría desde el inicio; g) elaboración de la versión final del Umbral Superior de Calidad Democrática.

12 En cada condición de evaluación se establecieron 3 situaciones: 0 puntos = incumplimiento severo; 1 punto = cumplimiento parcial; 2 puntos: pleno cumplimiento. Por ejemplo, en la condición "percepciones ciudadanas sobre la diversidad de partidos políticos" (información proveniente de la encuesta nacional de UNIMER) se definió la siguiente situación:

Puntos	Situación	Definición
2	Pleno cumplimiento	Más del 70% de las personas encuestadas considera que en Costa Rica hay suficiente variedad de partidos políticos para elegir en los procesos electorales (mayoría absoluta).
1	Cumplimiento parcial	Entre el 30% y el 70% de las personas encuestadas considera que en Costa Rica hay suficiente variedad de partidos políticos para elegir en los procesos electorales (mayoría absoluta).
0	Incumplimiento severo	Menos del 30% de las personas encuestadas considera que en Costa Rica hay suficiente variedad de partidos políticos para elegir en los procesos electorales (minoría pequeña).

Esta era una de las dos condiciones de evaluación del indicador "alta estima ciudadana por el pluralismo político". El puntaje que podía obtenerse oscilaba entre 0 puntos (las dos condiciones se incumplen severamente) y 4 puntos (las dos condiciones se cumplen plenamente). La expresión matemática es la siguiente: el puntaje total obtenido por un indicador es $P_i = \sum p_i$, donde p_i es el puntaje obtenido en la condición "ci" y P_i es el puntaje total obtenido en el indicador i .

Para cada indicador se creó una escala discreta, cuyo rango es $0 < P_i = (p_{max} * C)$, donde C es el número de condiciones de evaluación, p_{max} es el puntaje máximo que puede obtenerse en una condición (2 puntos). Para cada indicador se estandarizó una escala, con el fin de ubicar el puntaje obtenido dentro de un rango continuo cuyos extremos son 0 (incumplimiento severo en todas las condiciones) y 1 (cumplimiento pleno en todas las condiciones). Así: $P_i = (P_i / p_{max}) * 100$, donde P_i es el puntaje total relativo (comparado con el puntaje máximo posible) obtenido en el indicador i . En todos los casos, la transformación del puntaje obtenido en la escala ordinal de cinco categorías se cionó a los siguientes criterios:

Escala	Categoría	Puntaje relativo Pri
5	Cumplimiento pleno	0%
4	Cumplimiento general	1- 33,3%
3	Cumplimiento medio	33,4% - 66,6%
2	Incumplimiento general	66,7% - 99,9%
1	Incumplimiento severo o frontera crítica	100%

13 Los resultados prácticos de la aplicación de la fórmula de agregación del juicio compartido y del juicio de expertos se presenta en la siguiente tabla:

Resultado del juicio compartido	Resultado del juicio de experto	Calificación final
Cumplimiento medio (3)	Cumplimiento medio (3)	Cumplimiento medio (3)
Cumplimiento medio (3)	Cumplimiento general (4)	Cumplimiento medio (3)
	Incumplimiento general (2)	Cumplimiento medio (3)
Cumplimiento medio (3)	Cumplimiento pleno (5)	Cumplimiento general (4)
	Incumplimiento severo (1)	Incumplimiento general (2)
Incumplimiento general (2)	Cumplimiento pleno (5)	Cumplimiento medio (3)
Cumplimiento general (4)	Incumplimiento severo (1)	Cumplimiento medio (3)
Incumplimiento severo (1)	Cumplimiento pleno (5)	Cumplimiento medio (3)
Cumplimiento pleno (5)	Incumplimiento severo (1)	Cumplimiento medio (3)

14 Ello ocurrió en el dominio sobre calidad democrática de la competencia electoral. Cabe señalar, sin embargo, que la colaboración recibida había sido de carácter puntual. Hubo casos (3) en que los investigadores participaron en la evaluación, pero en temas distintos a los de su contribución. En todo caso, lo ideal sería que ninguna de estas situaciones se diera.