**Tratamiento de Variables FONDECYT 1170239**

Este documento en formato Word busca dar una narrativa de la recodificación y construcción de variables. Principalmente en caso de que alguien requiera transparentar los criterios que se eligieron en base a los datos, o las decisiones respecto a la información contenida.

Cabe indicar que algunas observaciones se suprimieron a partir de Octubre de 2018, por poseer una ocupación incompatible:

* 3061358
* 3137017
* 3146980
* 3147305
* 3149284

Las variables resaltadas en amarillo, indican actualizaciones del documento en la definición de las mismas.

La variable Estrechez Económica **(estrechez)**, se obtuvo a partir de la pregunta A61 (“61. Por favor, indique la frase que más se acerca a la realidad de su hogar: “Pensando en el total de ingresos mensuales de su hogar, este…”). Se recodifica “1”, para aquellos casos que responden 3 (“No les alcanza, tienen algunas dificultades”) o 4 (“No les alcanza, tienen grandes dificultades”) en ambos ítems y "2" para aquellos que responden 1 ("Les alcanza bien, pudiendo ahorrar") y 2 ("Les alcanza justo, sin grandes dificultades").

El ítem A52 Sexo, se recodifica en una nueva variable **(mujer),** como mujer (2)=1, y hombre (1)=0. Además se incorpora otra variable (**hombre**), como hombre (1)=1.

La variable Síntomas Depresivos **(depre)** se construyó en base a las preguntas A48 y A49, y se recodifica como 1, para aquellas observaciones que presentan “Sí” (1) en ambas preguntas (2 síntomas depresivos).

La pregunta A21 se denominará **“autonaq”,** homologable a la autodefinición de violencia (“Si definimos el acoso como una situación en la que una persona percibe que continuamente es objeto de conductas negativas frente a las cuales tiene dificultades para defenderse ¿según esta definición, usted ha sido acosado/a durante los últimos 6 meses?”).

Con el ítem A21.1 “Y, ¿qué tan frecuentemente ha ocurrido ese acoso?") se creó una nueva variable llamada **frec\_acos21\_1** (etiqueta "frecuencia con autonac") y se recodifica en "1" en aquellos casos cuya frecuencia comprenda entre 3 (“Varias veces por semana”) y 4 (“Prácticamente a diario”), mientras que se recodifica como "0" las categorías de respuesta entre 1 ("pocas veces) y 2 ("De vez en cuando").

Se creó la variable Consumo de Psicotrópicos **(“psicotrop”),** en base a las preguntas “A46\_top\_1”, “A46\_top\_12” y “A46\_top\_13”. Se considera consumo de psicotrópicos si el encuestado consumió al menos 1 tipo de psicotrópico lo que fue codificado con "1" y con "0" el no consumo.

La variable cuya etiqueta es atención de público **("atpub")**, hace alusión a la pregunta A17 “En su trabajo, ¿tiene habitualmente contacto directo con público, de viva voz o por teléfono (pacientes, clientes, alumnos, apoderados)?” (1= “Sí” y 0="No").

La suma de las preguntas A 47 del 1 a la 6, ("A47\_top\_1", "A47\_top\_2", "A47\_top\_3", "A47\_top\_4", "A47\_top\_5" y "A47\_top\_6") se codifica con nombre “**k6**” (para esta pregunta los 99 no son recodificados). Se registran algunos casos que poseen datos perdidos (99=.) . Dado que para calcular Distrés Psicológico K6 se requiere la suma de los puntajes para cada observación, puede haber respondientes que por el hecho de haber contestado en categorías 3 (“La mayoría de las veces”) o 4 (“Siempre”) en algunos de los 6 ítems, pero perdidos en el resto, la prevalencia de Distrés sea subestimada (similar a un 0, ya que las no-respuestas, al no estar asociadas a un número, no adhieren puntaje a la suma de los puntajes brutos).

Habiendo hecho un análisis de cada encuestado que no-respondió algún ítem de los pertenecientes a k6, se acordó que una eventual imputación de los datos no afectará la categoría en la que resultarán recodificados cada respondiente, por lo que no pareció necesaria una imputación de los datos para esta variable.

Se crea una variable **"k6\_2"** en base a la suma de cada uno de los 6 ítems, pero con los puntajes 99 ya recodificados como perdidos (99=.).

Se crea la variable **"k2\_muy elevado"** para los puntajes mayores a 12 en la variable k6\_2. Adicionalmente, se crea la variable **"k2\_elevado"** para los puntajes mayores a 6.

Las preguntas del AUDIT Corto **("auditc")** concerniente a los ítems A51\_1, A51\_2 y A51\_3, al no poseer como alternativa de respuesta 99’s (NR’s), se calcula mediante la sumatoria de los puntajes (que van de 0 a 12). Se recodifica como consumo riesgoso (1") si presenta 3 o más puntos para mujeres, y 4 o más para hombres.

A22 (“"Si definimos el acoso como una situación en la que una persona percibe que continuamente es objeto de conductas negativas frente a las cuales tiene dificultades para defenderse ¿según esta definición, usted ha sido acosado/a durante los últimos 6 meses"), se le etiqueta como “**presencia acoso psicologico ecqotesst**”. Por su parte, la pregunta A23 ("¿Con qué frecuencia le ha ocurrido esto?"), se etiqueta **“frec\_acos\_ecotest”.**

Se generó la variable **agr\_fisica** a partir de la variable A24 (“¿En su trabajo principal, en los últimos 12 meses, ha sido víctima de agresión física?”), descartando los valores 99’s.

También se generó la variable **acos\_sex** a partir de la variable A26 (“En los últimos 12 meses, en su trabajo principal ¿ha sido objeto de palabras o gestos de carácter sexual no deseado?”), descartando los valores 99’s.

El ítem A28, “**testigo**” (bystander) se etiqueta “testigo de acoso sexual” ("¿En los últimos 12 meses, en su trabajo principal ¿usted ha sido testigo de conductas, palabras o gestos de carácter sexual no deseado?").

La variable **acos\_psico\_freq** se genera en base a la pregunta A23 (“¿Con qué frecuencia le ha ocurrido esto?”), seleccionando sólo aquellos casos que indican una frecuencia “seguido” (2) y “muy seguido” (3).

Las variables **jerarquia\_acoso** y **jerarquia\_acoso2** se obtuvieron de las preguntas A29\_01 y A29\_02 (Si usted ha sido objeto de acoso psicológico, agresión física o palabras o gestos de carácter sexual no deseados ¿Quiénes han sido principalmente los agresores?), eliminando las categorías “nr”(99’s).

La variable **gse\_ac3** resultará codificada de “gse” (variable proveniente de categorización realizada por Adimark), El Grupo Socioeconómico de pertenencia del(a) encuestado(a) fue calculado para dar cuenta de los grupos Socioeconómicos D (4) y E (5) (Ansoleaga, Díaz &amp; Mauro, 2016b). La presente clasificación fue elaborada por la empresa Adimark (actual GFK) en base a la información entregada por el CENSO aplicado el año 2002 y que utiliza entre sus indicadores, el tramo de ingresos mensuales del hogar y el nivel de estudios del jefe de hogar (Adimark, 2004).

La variable “Supervisa a otros” (**superv**) se generó a partir de A2 (“En la empresa u organización donde usted trabaja ¿Supervisa a otros/as empleados/as?“), en la que valores “nr” (99’s) fueron recodificados como valores perdidos.

La variable **SECT\_PRIV** se generó a partir de la variable A4 (¿Usted trabaja para un organismo público o para una empresa privada?), indicando aquellos casos que pertenecen al sector privado (2=1), en comparación al “sector público” (1=0), “De las fuerzas armadas y de orden (civiles)” (3=0) y “Otros” (4=0). 88’s y 99’s (No sabe o No Responde) fueron recodificados como valores perdidos.

La variable Ausencia de Contrato de trabajo (**cont\_trab**), se generó a partir de la variable A5 (“En este trabajo: Tiene contrato o acuerdo de trabajo escrito?”), siendo “No” (0=1) la categoría de interés, en comparación a “Sí” (1=0). “No Sabe” (88), “No Responde” (99) fueron considerados valores perdidos.

La variable **jef\_hogar** proviene de la pregunta A55 (“¿Es usted el/la jefe/a de hogar?”), descartando los valores 99’s como valores perdidos.

Se genera la variable “Ausencia de Organización Sindical en su Empresa/Org” (**org\_sind**)proviene de la preguntaA40 (“¿En su empresa o institución existe alguna organización sindical o gremial?”), en la que recodifica por “No” (0=1), mientras que quienes responde “Sí” (1=0), descartando “No Sabe” (88=.) y “No Responde” (99=.) al tratarlos como valores perdidos.

La variableContrato Definido **(cont\_def)** proviene de la pregunta A6 (“¿Qué tipo de contrato o acuerdo de trabajo tiene?”), siendo la variable de comparación el contrato “Indefinido” (1=0), respecto a otros tipos de contrato formales e informales (2 a 9=1). Las categorías “No sabe” (88) y “nr” (99) fueron codificadas como valores perdidos.

**Educacion\_media**, de la pregunta A54 (“Indique su nivel educacional”), en la que se colapsan las categorías Menores a la Obtención de Educación Media Completa 1 a 4 (“No ha asistido a educación escolar” a Educación básica completa”).

La variable Carga doméstica (**carga-dom**) proviene de la variable A16 (“Y sin considerar su trabajo remunerado, durante la semana, ¿usted dedica además tiempo para hacer tareas domésticas y/o de cuidado, pero no remunerado?”), descartándose los valores “nr” (99) como perdidos.

La variable **turnos** se obtuvo del ítem A10 ("Respecto a su horario de trabajo, ¿este es?"), y en la que las categorías de respuesta de relevancia serán aquellos trabajos distintos al “Solo diurno (de día)” (1=0) (2 a 11= 1).

La variable **zona** concerniente a la región administrativa, fue recodificada dependiendo de cada región, generando **Santiago**, **valpo**, y **concep**, para cada una.

**OCUPACIONES**

Una es la variable **A1x1\_Coded\_1,** recodificación realizada por GFK, en la que agrupan en 11 categorías las actividades de la pregunta A3. No obstante, no entregan mayores detalles respecto al criterio utilizado para recodificar la ocupación en base a las categorías presentadas.

De igual forma, también existe otra variable vinculada a ocupación (**ocux**) y posiblemente elaborada por GFK, aunque tampoco se manejan antecedentes del criterio utilizado para recodificar las categorías de esta forma.

Por último, también se constata una variable de nombre “**ocuprev**” con 9 categorías, que parecen ser similares a las presentadas por el Proyecto Araucaria (guarda relación con la diapositiva n°64 de la presentación disponible en el siguiente enlace: <http://www.proyectoaraucaria.cl/documentos/20120405085755calidad-empleo.ppt>). Esta variable fue elaborada por Amalia, Ximena y Magdalena a finales de 2018.

Nota. En algunas bases de datos la categoría 5 se encuentra etiquetada como “Empleados/as de Oficina”, siendo que debiese ser “Vendedores”. En tales casos, existirá la variable **ocuprev\_rec** que posee esta categoría recodificada.

De esta última se genera la variable **prev3**, estando compuesta por Trabajadores de los servicios (8), Vendedores (5), Empleados/as de Oficina (4), Profesionales (2) y Jefaturas Intermedias (1).

Posteriormente, para el artículo de Magdalena, se genera la variable **ocuprev\_nocalif,** a partir de la variable ocuprev\_rec, seleccionando a Operarios (6=1), Técnicos (3=1), Conductores de transporte (7=1) y Trabajadores no-calificados (9=1). Elaborado en el contexto del artículo para cultura organizacional. Es lo mismo que prev3, pero invertida.

**Precariedad**

Con la finalidad de construir la variable Precariedad, se generan algunos ítems y recodifican otros. En primera instancia, se recodifica **“prec1”,** proveniente del ítem A5 ("En este trabajo ¿Tiene contrato o acuerdo de trabajo escrito?"), de manera que la ausencia de un contrato indique la presencia de un componente de precariedad (0 “No” =1), mientras que las categorías de respuesta “no sabe” (88) y “no responde” (99) no serán consideradas.

Adicionalmente , se crea el indicador **“prec2”**, proveniente del ítem A6 ("¿Qué tipo de contrato o acuerdo de trabajo tiene?"), cuyas categorías de relevancia serán las siguientes: “Con plazo de término (temporada, obra o faena)” (2), “A contrata” (3), “Contrato a honorarios (boletea)”(4) y “Otro”(5)(2 a 5=1; 1=0 y 6 al 99=0).

Otro indicador que se incorpora es el de nombre **“prec3 ”**, que resulta de la categorización de aquellos respondientes que indiquen en el ítem A8 (“"Independiente de su contrato de trabajo, habitualmente en su trabajo principal, ¿cuántas horas trabaja en la semana?"), trabajar en “Jornada parcial: menos de 45 horas semanales (2), y que señalen en el ítem A9 ("¿Usted desearía trabajar?"), “más horas a la semana” (2). Sumado a la primera condición, se construye el indicador **“prec4”** que agrupa quienes responden trabajar más de 45 horas semanales (3) en el ítem A8, y al mismo tiempo señalan desear trabajar “Menos horas a la semana” (1) en el ítem A9.

Cabe señalar que posteriormente se crea **"prec3r"** en que se consideró sólo a las personas que desean trabajar menos horas a la semana (1) o más horas a la semana (2) y se recodifica 1 y 2= 1, mientras quienes se encuentran satisfechos con las horas que trabajan (3) se recodifica como 0.

El indicador **“prec5”** se crea a partir de la pregunta A10 ("Respecto a su horario de trabajo, ¿este es?"), y en la que las categorías de respuesta de relevancia serán aquellos trabajos distintos al “Solo diurno (de día)” (1=0) (2 a 11= 1).

Un indicador relevante, denominado “**prec6**”, se elabora en base al ítem A11 ("Podría decir si la remuneración de su trabajo principal es:") y en la que las categorías de respuesta relevantes son aquellas que son distintas a “Fija” (1=0; 2 al 88= 1).

Por otro lado, **"prec7"** se refiere a estabilidad en el empleo, se elabora en base al item A34.7 ("Mi estabilidad laboral es baja"), 1 (muy de acuerdo) y 2 (de acuerdo) se recodifica como "1", mientras que 3 (en desacuerdo) y 4 (muy en desacuerdo), se recodifica como "0".

La variable **"prec8"** se refiere al desempleo, obtenida a partir de la pregunta A15 ("En los últimos 12 meses ¿usted ha estado desempleado?"), 1=Si; 0=No.

La variable **"prec9"** se refiere al ingreso personal obtenida a partir de la pregunta A58 ("¿Nos podría decir en cuál de los siguientes tramos se encuentra su ingreso personal mensual?) se recodifica1 ("Menos de $150.000 mensuales") y 2 ("De $150.001 a $350.000 mensuales")= 1 y el resto de las categorías que obedecen a tramos de mayores ingresos se recodifica como 0.

Una vez elaboradas las variables relativas a precariedad se genera la variable "precariedad" **(prec),** producto de la suma de las variables desde prec1 a prec9, sin considerar prec3 ni prec4 (prec1 + prec2 + prec3r + prec5 + prec6 + prec7 + prec8 + prec9). Posteriormente, la variable se dicotomizó como "1" para aquellos encuestados que puntúen más de 2 puntos y "0" para el resto (0 a 2=0 ; 3-7=1).

**precr1**= Índice de precariedad incluyendo vulnerabilidad, producto de la suma de prec1, 2, 3, 5, 6, 7, 8 y vul. No se incluye la variable de prec9. Puntajes mayores o iguales a 3 indican precariedad.

También se aborda la variable **Antigüedad en el cargo** en base al ítem A13 (“¿Puede decirnos cuánto tiempo lleva trabajando en esta empresa u organización?"), no considerando a aquellos que poseen menos de 6 meses (1) (2 y 3 =1; NR(99) =.).

La pregunta A14, relativa al Politrabajo, denominada **“politrab”** ("Además del trabajo principal, ¿usted destina tiempo a otro trabajo pagado?"), categoriza a aquellos que “Sí” destinan tiempo a otro trabajo pagado (1 =1; 99=.)

Con la finalidad de medir la variable Desempleo, etiquetada como **“desempleo”,** se ocupará la pregunta A15 ("En los últimos 12 meses ¿usted ha estado desempleado?" ) se considerará a aquellos que hayan respondido afirmativamente (1=1), y se categorizando aquellos casos NR como perdidos (99=.).

La variable Trabajo doméstico, etiquetada como **“trabdom”,** utilizará la pregunta A16 ("Y sin considerar su trabajo remunerado, durante la semana, ¿usted dedica además tiempo para hacer tareas domésticas y/o de cuidado, pero no remunerado?"), siendo “Sí” (1) la categoría de respuesta relevante, excluyendo NR’s (99=.).

A partir del artículo a publicar por Amalia y Ximena, se recodifican las siguientes variables:

**Educacion277**, de la pregunta A54 (“Indique su nivel educacional”), en la que se colapsan las categorías 1 a 3 (“No ha asistido a educación escolar” a Educación básica completa”) en "Sin Educacion o Basica"(1), 4 a 6 (“Educación Media Incompleta” a “Educación Técnica Incompleta”) en "Media o Tecnica Incomp"(2), 7 a 8 (“Técnica Completa” a “Universidad Incompleta”) en "Tecnica Comp o Univ Incomp"(3) y 9 a 10 (“Universidad Completa” a “Estudios de Postgrado”) en "Universitaria Comp o Postgr" (4).

**ocuprev\_xi\_am277** se genera a partir de la categorización de ocupaciones utilizado para el proyecto Araucaria (ocuprev), colapsándose en 5 categorías: Jefaturas Intermedias (1) (1=1), “Profesionales y Técnicos” (2) (2/3=2), “Empleados Oficina y Vendedores”(3) (4/5=3), “Operarios, conductores y trabajadores calificados” (6, 7 y 8=4), y “No calificados” (5) (9=5).

**ing\_hog\_xi\_am277** seobtiene a partir de una pregunta por los tramos de ingresos mensuales (A62), la que es colapsada en “Menos de 540.000 pesos”(1) (1 a 3=1), “Desde 540.001 hasta 913.000 pesos"(2) (4=2), “Desde 913.001 o más” (3) (5=“De 913.001 a 1.567.000” y 6=“Más de 1.567.001”). Las respuestas “No sabe” y “NR”’s se consideraron como valores perdidos.

**Job Content Questionnaire (JCQ)**

Se crea en base a la pregunta 18, las siguientes variables:

**“Khab”** (ítems A18\_top\_1 a 3) referente a desarrollo de habilidades.

**“Kcont”** (ítems A18\_top\_4 a 5), referente a control.

Las dos variables anteriores se sumaron y se creó la variable **"latdec", referente a latitud decisional.**

En base a la variable "latdec", se crea la variable **“latdec\_dic"** dicotomizándola en aquellos encuestados que presentan puntaje de 0 a 13=1 (Baja latitud) y de 14 a 20 (Alta latitud), considerando que “se sugiere utilizar como punto de corte el percentil 25 = 13 puntos, ya que es en este cuartil expuesto en el que se observan diferencias en la probabilidad de riesgo (Ansoleaga, 2013, p. 148).

Se combinan Demandas Psicológicas (JCQ) y Demandas Emocionales (ISTAS 21), en **“democ”** (ítems A18\_top\_ 6 a 10 y A19\_top\_1 a 4). La categoría de respuesta “siempre” de los ítems de demandas emocionales se asumen como “muy de acuerdo” de demandas psicológicas de Karasek, ya que ambos tienen la misma cantidad de categorías de respuesta.

Se dictomiza la variable "democ", en una variable llamada **"democ\_dic"** Hasta el 20 se considerará baja. De 0 a 20 puntos igual a “Baja” (0), de 21 a 36 igual a “alta” (1).

Se crea la variable **"aposoc\_tot",** referente al apoyo social, correspondiente al ítem A35 ("A35\_top\_1" "A35\_top\_2" "A35\_top\_3" "A35\_top\_4" "A35\_top\_5" "A35\_top\_6") producto de la suma de los puntajes, se codifica de 6 a 17 igual 1 (bajo apoyo social) y de 17 a 24 igual a 0 (alto apoyo social).

Se crea la variable **"tenpsi",** referente a tensión psíquica, correspondiente a la suma de las variables dicotomizadas de latitud decisional **("latdec\_dic")** y demandas emocionales ("democ-dic"). Una vez creada la variable se recodifica en 1 para aquellos respondientes que tienen Baja latitud decisional (1 en "latdec\_dic") y altas demandas emocionales (1 en "democ\_dic") (de 0 a 1=0 y 2=1).

Se crea la varialble **"iso"** que alude a la variable latente ISOSTRAIN, correspondiente a la suma de las variables dicotomizadas **"tenpsi"** (tensión psíquica) y **"aposoc\_tot"** (bajo apoyo social), donde 1 corresponde a la presencia de ambos atributos (2=1 y 0 a 1=0).

Se crea la variable **"salgen"** que alude a salud general (A44 "En general, diría usted que su salud es…"), se dicotomizó, de 1 a 3 (Excelente, muy buena o buena) igual 0, mientras que 4 y 5 (Mala y muy mala) igual 1.

Se crea la variable **"climasex",** correspondiente a la pregunta A36 "¿En su lugar de trabajo hay un clima sexista?". Se consideró un valor perdido la categoría de respuesta 88 (No sabe).

Se crea la variable **"culorg1 ",** correspondiente a cultura organizacional medida en la pregunta A 37\_1 ("En mi empresa, (institución, organización) lo importante es cómo se hacen las cosas, no importan tanto los resultados. v/s En mi empresa, (institución, organización) lo importante son los resultados, no importa tanto cómo se hacen las cosas."). Al igual que los dos siguientes (**"culorg2"** y **"culorg3"**), al ser una variable bipolar se dicotomiza aquellas categorías de respuesta que están "más asociadas" a un polo o el otro, agrupando en el de la derecha (1) del 4 al 6 y en el de la izquierda (0) del 1 al 3. (1,2 y 3=0; 4, 5 y 6=1).

En base a la suma de las tres variables antes mencionadas ("culorg1", "culorg2" y "culorg3") se genera la variable **"cultura",** que se dicotomiza de acuerdo a las personas que señalan rabajar en una organización con cultura más asociada a los resultados, al cumplimiento de tareas y al control (4 a 6=1 y 1 a 3=0). Por otra parte, **culorg1\_ext** a **culorg3\_ext**, sólo considerará los extremos (los valores 3 y 4 se igualarán a 0, el resto =1). En cambio, **proc\_res**, **bienest\_tarea** y **laxo\_estr** reconocen los valores balanceados (3 a 4 =1, el resto= 0). Posteriormente, se generaron una serie de variables relacionadas con las clases latentes (valores a priori, a posteriori, entre otras), de los que resultan de particular importancia las variables **rec\_expclass\_CULORG\_1 a 6**  o **rec\_expclass\_CULORG\_alt\_1 a 6** (dependiendo de los parámetros especificados en el modelamiento de clases latentes)que definen la clase a la que debiese pertenecer cada caso si obtiene un valor a-posterior mayor a 0,5 en alguna de las clases latentes. Finalmente, la variable “**cluster\_rec2**” y “**cluster\_rec2\_alt**” (dependiendo de los parámetros especificados en cada modelo de clases latentes).

**DER**

La variable **esfuerzo** se obtuvo de la suma de los ítems A34\_top\_1, A34\_top\_2 y A34\_top\_3, una vez fueron transformadas en e1 a e3 una vez se les eliminó la categoría “nr” (99). De igual forma, la variable **recompensa** se obtuvo de la suma de puntajes de A34\_top\_4 a A34\_top\_10, una vez fueron transformadas en r1 a r7 una vez se les eliminó la categoría “nr” (99). La variable **bajarecompensa** y **altoesfuerzo** se obtuvieron dicotomizando estas variables en valores menores a la mediana, siendo 0 a 17 para baja recompensa, y valores mayores o iguales a 8 para alto esfuerzo. El **desbalance** resultó de la división del puntaje de esfuerzo por el puntaje de recompensa multiplicado por 0,4285 (esfuerzo / (recompensa \* 0.4285)). Valores mayores o iguales a 1 se consideraron en desbalance (**desbalance\_dic**).

**NAQ-R**

Para cada una de las variables relativas al cuestionario de conductas negativas NAQ-R (A20\_1\_Top\_ 1 a 12, y A20.2\_Top\_13 a 22), se generan las variables “NAQ1” al “NAQ22”. Posteriormente, de las mismas 22 variables, se crean otras dicotomizadas de acuerdo a criterio operacional OCM ("3. al menos 1 semana" o "4. a diario"), cuyo nombre va de “**NAQ1\_rec” a “NAQ22\_rec”** (la recodificación consistió en tratar a las categorías 0, 1 y 2 como “0”, mientras que 3 y 4 como “1”). Una vez se obtienen dichas variables se genera otra, producto de la suma de cada una de estas variables recodificadas (**NAQ\_Dic**). De ella, se genera "**NAQ\_Dic\_logit**", la que categoriza la suma de puntajes dicotomizados (a aquellos que reportan al menos 1 semana en cada ítem) en 0, si no registran conductas con una frecuencia mayor o igual a 1 semana, 1, si poseen una conducta con una frecuencia igual o mayor a una semana, y 2, si reportan 2 o más. Esta variable buscó encontrar diferencias entre quienes sólo reportan una conducta respecto a quienes reportaban 2 o más. La variable NAQ\_Dic nos permite generar los criterios de clasificación operacional por cada observación: el indicado por Leymann (**NAQ\_Leyman**), que alude a aquellos que en NAQ\_Dic hayan obtenido puntajes mayores a 0 (ya que requiere que en al menos un ítem indique una frecuencia de al menos 1 semana), siendo los puntajes de 1 a 22 =1 (0=0); el indicado por Mikkelsen y Einarsen (**NAQ\_MyE**), que considera aquellos respondientes que en NAQ\_Dic hayan obtenido puntajes mayores a 1, siendo los puntajes de 2 a 22= 1 (entre 0 y 1= 0); por último, el criterio de Agervold (NAQ\_A), requiere que los respondientes indiquen en al menos 3 ítems una frecuencia mayor a al menos una semana (3 a 22 =1;0 a 2=0).

Se generó la variable "**NAQ\_sum**" de la suma de puntajes brutos de la prueba NAQ (de los ítems 1 al 22). A partir de esta última variable, se subdividió conforme a su distribución en cuartiles, "**NAQ\_sum\_cuartil**", siendo los puntajes 0 el primer cuartil (24% ponderado aprox.), de 1 a 4 el segundo cuartil (30%), de 5 a 10 el tercero (23%) y el cuarto, de 11 hasta 88(24).

También se incorporaron las variables con la suma de los ítems respondidos "Al menos 1 vez a la semana" y "A Diario" (lo que Notelaers & Einarsen llaman "dichotomized sum score approach"). Estos son: **NAQ\_R\_AF\_rec**, **NAQ\_R\_AP\_rec** y **NAQ\_R\_RT\_rec**; luego estos ítems se recodificaron de acuerdo a cada uno de los criterios (al menos 1 ítem, por Leymann, al menos 2, por Mikkelsen & Einarsen, y al menos 3, por Agervold): **NAQ\_AF\_Leyman**, **NAQ\_AP\_Leyman**, **NAQ\_RT\_Leyman**, **NAQ\_AF\_MyE**, **NAQ\_AP\_MyE**, **NAQ\_RT\_MyE**, **NAQ\_AF\_A**, **NAQ\_AP\_A** y **NAQ\_RT\_A**.

**Curvas ROC NAQ-R**

De la suma de puntajes brutos del NAQ, se agregó la variable "**ARAUCARIA\_ACOSO**" que suma las respuestas a las preguntas A22, A24 y A26 (sexual, psicológico y físico, tratando los perdidos como 0’s), disponibles sin datos perdidos en acos\_sex, A22 y agr\_fisica presentadas en el marco de la Encuesta Araucaria. Aquellos casos que obtengan 2 o más puntajes, serán considerados como expuestas a acoso (2 o más=1). De la combinación de acos\_sex, A22 y agr\_fisica con la de acoso laboral del NAQ-R (autonaq), se generó una variable amplia de categorización de acoso laboral ("**ACOSO\_AUTO**") que agrupa a estas 4 preguntas que aluden a victimización percibida de acoso hacia su persona tanto en los últimos 12 meses como en los últimos 6 *(autocategorización bullying*, acoso psicológico, físico, y sexual). Finalmente, esta variable, al ser un criterio menos estricto que el anterior, se determinó en la tesis de Andrés González a determinar alguna forma de incivismo o acoso de menor intensidad, se recodifica en aquellos que contestaron a al menos alguna forma de victimización (1 o más=1) ("**ACOSO\_AUTO\_REC**"). Se categorizó a **NAQ\_sum\_ROC2** como el punto de corte para diferenciar (9 a 88= 1, 0 a 8=0). De igual forma, "**NAQ\_sum\_ROC**", producto del punto de corte únicamente en base al criterio de autocategorización del NAQ (de 0 a 11, se recodifica como 0, mientras que de 12 a 88, 1). "**NAQ\_sum\_ROC3\_cat**" que combina ambas variables y que agrupa 3 categorías: no expuesto (puntajes de 0 a 8), ocasionalmente/ampliamente expuesto (puntajes de 9 a 11), y expuesto (puntajes de 12 a 88).

Nota: Cabe señalar que los puntos de corte se construyeron en base a medidas realizadas en el software R para el contexto de la tesis de Andrés González.

**Liderazgo Destructivo**

Para cada una de las variables de la Escala de Liderazgo Destructivo (A38\_top\_1 a 14), se generan las variables **LID\_DES1 a LID\_DES14**. Posteriormente, se dicotomizan las variables antes mencionadas de acuerdo al criterio de clasificación operacional OCM (esto es, si el encuestado responde "2. Bastante a menudo" o "3. casi siempre"), creando con ello las variables **LID\_DES1\_rec** a **LID\_DES14\_rec** (0 a 1 = 0; 2 a 3= 1). En base a estas variables se elabora cada una de las formas de liderazgo producto de la suma de puntajes para cada una de las variables de cada subescala: para liderazgo Laissez-Faire (“**LF\_rec”**) se suman las variables LID\_DES3\_rec, LID\_DES5\_rec, LID\_DES7\_rec y LID\_DES9\_rec; para liderazgo Autoritario (“**LA\_rec”**) se suman las variables LID\_DES1\_rec, LID\_DES4\_rec, LID\_DES8\_rec y LID\_DES10\_rec; por último, para liderazgo Constructivo (“**LC\_rec”**) se suman las variables LID\_DES2\_rec, LID\_DES6\_rec, LID\_DES11\_rec, LID\_DES12\_rec, LID\_DES13\_rec y LID\_DES14\_rec. En última instancia, se generan las variables dicotomizadas de cada una de las formas de liderazgo de acuerdo al criterio Leyman (**LF\_Leyman, LA\_ Leyman y LC\_Leyman**) que alude a aquellos que en cada una de las formas de liderazgo (LF\_rec, LA\_rec yLC\_rec, respectivamente) hayan obtenido puntajes mayores a 0 (ya que requiere que en al menos un ítem indique una frecuencia de al menos Bastante a menudo), siendo los puntajes de 1 o más=1 (mientras que 0=0).

Nota: Una vez se descartan algunos ítems producto del análisis factorial, las variables LC\_rec y LF\_rec, así como LC\_Leyman y LF\_Leyman, son recodificadas de acuerdo a los ítems resultantes de la validación, descartando los ítems 2,3 y 13.

De igual forma, se incorporan las variables "**LA\_rec\_logit**", "**LF\_rec\_logit**", "**LC\_rec\_logit**" (invertido, se buscó la ausencia de liderazgo constructivo) y "**NAQ\_Dic\_logit**", las cuales categorizan la suma de puntajes dicotomizados (a aquellos que reportan al menos 1 semana en cada ítem) en 0, si no registran conductas con una frecuencia mayor o igual a 1 semana, 1, si poseen una conducta con una frecuencia igual o mayor a una semana, y 2, si reportan 2 o más. Esta variable buscó encontrar diferencias entre quienes sólo reportan una conducta respecto a quienes reportaban 2 o más.

Posteriormente, se genera la suma de las variables dicotomizadas de Liderazgo Destructivo, con aquellos ítems que resultaron de la validación (5,7,9,1,4,8 y 10). Se sumaron las frecuencias de aquellos ítems en los que se responde al menos Bastante a menudo (es decir, con valores 2 o 3). Se entiende como la suma de los puntajes dicotomizados de acuerdo a criterio de clasificación operacional. Valores iguales o mayores a 4 quedaron definidos como 4 o más en la variable **LID\_DES\_Dic**. La variable posteriormente es colapsada en 3 categorías: ausencia de conductas (0), una conducta (1) y 2 o más conductas (2) en **LID\_DES\_Dic\_logit**. Lo mismo ocurre con las variables "**LA\_rec\_logit**", "**LF\_rec\_logit**", "**LC\_rec\_logit**" (invertido, se buscó la ausencia de liderazgo constructivo) y "**NAQ\_Dic\_logit**", las cuales categorizan la suma de puntajes dicotomizados (a aquellos que reportan al menos 1 semana en cada ítem).

**Vulnerabilidad**

Se crea la variable “Escala de Vulnerabilidad” **(“vul1” a “vul7”)**, correspondiente a la pregunta A 39 “A continuación, queremos preguntarle en qué medida está de acuerdo con las siguientes afirmaciones sobre su trabajo” (A39\_top\_1 a A39\_top\_7). Cada uno de ellos se dicotomizó, siendo 3 y 4 igual a 1 (“Muy de acuerdo” y “de acuerdo”), mientras que 1 y 2 (“En desacuerdo” y “muy en desacuerdo”) igual 0.

Con la suma de las respuestas de las variables anteriormente mencionadas, se crea “Índice de Vulnerabilidad” **(“vul”)** yse recodificó de 0 a 2 igual 0 (ausencia de vulnerabilidad), y 3 a 7 igual 1 (presencia de vulnerabilidad).

**Satisfacción Laboral**

Se generaron las variables **s1 a s5**, provenientes de A43\_top\_1 a A43\_top\_5, pero recodificando la no-respuesta (“nr”) como valores perdidos (99=.). Estas variables se sumaron (valores perdidos se trataron como 0’s), resultando en la variable **satis**. La variable **satlab\_rec**, por su parte, alude a los casos que indicaron puntajes menores a la mediana (0 a 13 = 1, y 0 para los valores de 14 a 20).

**Consideraciones**

Para todas las variables en que no se mencione el tratamiento de datos perdidos, comprobar que las respuestas No Responde (“nr”o 99’s) se hayan codificado como valores perdidos.