An Exploratory Technique for Finding the Q-matrix in Cognitive Diagnostic Assessment: Combining Theory with Data

Catherine N. Close
Renaissance Learning
901 Deming Way Suite 301
Madison WI 53717

Mark L. Davison
University of Minnesota
182 Education Sciences Building
56 East River Road
Minneapolis, MN 55455
MLD@umn.edu

Ernest C. Davenport, Jr.
University of Minnesota
172 Education Sciences Building
56 East River Road
Minneapolis, MN 55455
LQR6576@umn.edu

1. Introduction

There is a growing interest in testing students formatively during the course of teaching in order to determine skills that students do or do not possess. Information about skill mastery or nonmastery can help tailor instruction to needs of the students. A useful tool for formative assessment is cognitive diagnostic assessment (CDA). CDA measures the specific knowledge structures and the processing skills that examinees possess to provide information about the cognitive strengths and weaknesses of examinees (Leighton & Gierl, 2007). The cognitive diagnostic models (CDMs) that CDA uses to relate the latent skills to observed behavior (tasks) require a Q-matrix having elements q_{jk} for J items and K attributes (Tatsuoka, 1983).

The Q-matrix embodies the design of the assessment instrument in use and in essence determines the quality of the resulting diagnostic information (Rupp and Templin, 2008). Therefore, test developers should ensure that the necessary procedures and expertise in cognitive theory are used in determining the Q-matrix for a diagnostic test. In this study, we propose components analysis as an exploratory technique for finding the Q-matrix in addition to using theory. Studies such as Templin and Henson (2006); and Henson, Templin, & Douglas (2007) have discussed factor analysis for cognitive diagnosis but not for Q-matrix development. Liu, Douglas & Henson (2009) address the use of factor analysis for Q-matrix development as discussed in Henson and Templin (2007), indicating that factor analysis can give a reasonable solution when the Q-matrix is not too complex. One study has examined a method for validating the Q-matrix (de la Torre, 2008) and so more studies are needed.

As such, the purpose of this study was to investigate components analysis, an exploratory technique that can be used to find the Q-matrix of a cognitive diagnostic test (when the number of skills is unknown) in data that satisfy the DINA (Deterministic Input Noisy "And") model of Haertel (1989). Extension of the proposed analysis to other conjunctive models such as the reduced version of the RUM model of Hartz (2002) is straightforward.

2. Method

Data were simulated using the DINA model (see the section on cognitive diagnostic model below for details). Because the Q-matrix is considered fixed, the Q-matrix was designed to measure 3 skills with 21 items. At least 2^K -

I items are required to measure *K* skills. Each skill level combination in our Q-matrix was measured by at least four items but no items measured all three skills. The item parameters, s_j and g_j, were simulated from a random uniform distribution, $s_j \sim U(.02, .05)$ and $g_j \sim U(.05, .25)$ Examinees' latent ability vectors, $\underline{\alpha}_i$, were simulated from a probit model having underlying latent variables that are multivariate normal with mean vector zero and correlations between the skills fixed at .50; $MVN(\underline{\mu}, \mathbf{\Sigma})$ where $\underline{\mu}$ is the mean vector and $\mathbf{\Sigma}$ is the correlation matrix. The proportion of the population assumed to have mastered the attributes, p_k, was fixed at 0.50. These parameter specifications are commonly used in published articles. The data were simulated in R, freely available software. In addition to simulated data real data was also analyzed. The real data used in the analysis were the fraction subtraction data collected by Dr. Kikumi Tatsuoka (Tatsuoka, 1983). The data were obtained from the Royal Statistical Society website, http://www.blackwellpublishing.com/rss/Volumes/Cv51p3.htm and are comprised of dichotomously-scored responses to 20 fraction subtraction test items from 536 middle school students.

2.1 Cognitive Diagnostic Model

Define X_{ij} as a binary indicator of whether examinee *i* performed task *j* correctly; q_{jk} as a binary indicator of whether attribute *k* is relevant for task *j*; and α_{ik} as a binary indicator of whether examinee *i* possesses attribute *k*. Then, the DINA model has the item response function (IRF),

$$p(X_{ij}=1 \mid \underline{\alpha}, s, g) = (1-s_j)^{\xi_{ij}} g_j^{1-\xi_{ij}} \tag{1}$$

where $\xi_{ij} = \prod_{k=1}^{K} \alpha_{ik}^{Q_{jk}}$, $s_j = p(X_{ij}=0 \mid \xi_{ij}=1)$, $g_j = p(X_{ij}=1 \mid \xi_{ij}=0)$

The DINA IRF can be rewritten into a principal components form. Due to the deterministic nature of the principal components analysis, the component form is the same as the factor model without the error term. Specifically, if Z_{ij} is the z-score of person *i* on variable *j*, the scalar form of the factor model is

$$Z_{ij} = \sum_m \lambda_{jm} f_{im} + e_{ij} \tag{2}$$

where λ_{jm} is the loading of item *j* on factor *m*, f_{im} is the factor score of person *i* on factor *m*, and e_{ij} is the error term. It follows that the expectation of Z_{ij} is the component form,

$$E(Z_{ij}) = \sum_m \lambda_{jm} f_{im} \tag{3}$$

where, λ_{jm} is the loading of item *j* on component *m*, f_{im} is the component score of person *i* on component *m*. The DINA IRF can be written in a similar manner,

$$p(X_{ij}=1 \mid \underline{\alpha}, s, g) = \sum_m \lambda^*_{jm} f^*_{im} \tag{4}$$

Asterisks indicate the parameters are in the raw score model (equation 4). Parameters in the standard score model are in equation 3 where the parameters are as defined previously. λ_{jm} equals 1 when the skill set corresponding to factor *m* is the minimal skill set sufficient for the solution of item *j*; else it is zero. If the data satisfy the DINA model, then $\lambda^*{}_{jm}$ equals $(1-s_j)$ when the skill set corresponding to factor *m* is the minimal skill set sufficient for the solution of item *j*; else it is zero. $f^*{}_{im}$ is either 1 or $(\frac{g_j}{1-s_j})$ depending on whether the examinee possesses the required skills for an item. With this reformulation,

$p(X_{ij}=1 \mid \underline{\alpha}, s, g) = \sum_m \lambda^*_{jm} f^*_{im} = \lambda^*_{jm'} f^*_{im'} = (1-s_j)\,(\frac{g_j}{1-s_j})$. The parameters in the raw score model (equation 4) can be related to those in the standard score model (equation 3) as follows: $\lambda_{jm} = \frac{\lambda^*_{jm}\hat{\sigma}^*_m}{\hat{\sigma}_j}$ and

$f_{im} = \frac{f^*_{im} - f^*_{.m}}{\hat{\sigma}^*_m}$ where, $\hat{\sigma}_j$ is the standard deviation for item *j*, $\hat{\sigma}^*_m$ is the standard deviation of the factor scores for a component m, and $f^{*}_{.m}$ is the mean component score for factor m. $\hat{\sigma}^{*}_{m}$ appears in both the numerator and the denominator of λ_{jm} and f_{im} because the variance of the components is 1 for standardized scores.

In equation 3, a component corresponds, not necessarily to a single skill, but a skill set and the items loading on the component are those for which the set is the minimally required combination of skills for the items. The components should follow a simple structure. As an example, if a test measures three skills, the skills corresponding to all combinations of the three skills are shown in Table 1. Using the DINA model, the resulting component scores can be represented as shown in Table 2. The probability of a correct response is then simply $\sum_f \lambda f$.

Table 1: Component Loadings

		Loadings (dichotomized)						
	Skill Combinations (Components)	Items Measuring Skill Set 1	Items Measuring Skill Set 2	Items Measuring Skill Set 3	Items measuring Skill Set4	Items Measuring Skill Set 5	Items Measuring Skill Set 6	Items measuring Skill Set 7
Item1	1	$1-s_1$	0	0	0	0	0	0
Item2	2	0	$1-s_2$	0	0	0	0	0
Item3	3	0	0	$1-s_3$	0	0	0	0
Item4	12	0	0	0	$1-s_4$	0	0	0
Item5	13	0	0	0	0	$1-s_5$	0	0
Item6	23	0	0	0	0	0	$1-s_6$	0
Item7	123	0	0	0	0	0	0	$1-s_7$

Table 2: Component Scores

		Component Scores						
	Skill Combinations Possessed	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 7
Person1	(1)	1	$\frac{g_2}{1-s_2}$	$\frac{g_3}{1-s_3}$	$\frac{g_4}{1-s_4}$	$\frac{g_5}{1-s_5}$	$\frac{g_6}{1-s_6}$	$\frac{g_7}{1-s_7}$
Person2	(2)	$\frac{g_1}{1-s_1}$	1	$\frac{g_3}{1-s_3}$	$\frac{g_4}{1-s_4}$	$\frac{g_5}{1-s_5}$	$\frac{g_6}{1-s_6}$	$\frac{g_7}{1-s_7}$
Person3	(3)	$\frac{g_1}{1-s_1}$	$\frac{g_2}{1-s_2}$	1	$\frac{g_4}{1-s_4}$	$\frac{g_5}{1-s_5}$	$\frac{g_6}{1-s_6}$	$\frac{g_7}{1-s_7}$
Person4	(1,2)	$\frac{g_1}{1-s_1}$	$\frac{g_2}{1-s_2}$	$\frac{g_3}{1-s_3}$	1	$\frac{g_5}{1-s_5}$	$\frac{g_6}{1-s_6}$	$\frac{g_7}{1-s_7}$
Person5	(1,3)	$\frac{g_1}{1-s_1}$	$\frac{g_2}{1-s_2}$	$\frac{g_3}{1-s_3}$	$\frac{g_4}{1-s_4}$	1	$\frac{g_6}{1-s_6}$	$\frac{g_7}{1-s_7}$
Person6	(2,3)	$\frac{g_1}{1-s_1}$	$\frac{g_2}{1-s_2}$	$\frac{g_3}{1-s_3}$	$\frac{g_4}{1-s_4}$	$\frac{g_5}{1-s_5}$	1	$\frac{g_7}{1-s_7}$
Person7	(1,2,3)	$\frac{g_1}{1-s_1}$	$\frac{g_2}{1-s_2}$	$\frac{g_3}{1-s_3}$	$\frac{g_4}{1-s_4}$	$\frac{g_5}{1-s_5}$	$\frac{g_6}{1-s_6}$	1

3. Analysis

Binary examinee responses from the simulated and real data were used to compute item correlations that were analyzed via principal components analysis (PCA) with promax rotation (kappa=4) to identify the skill sets. Components were extracted until every item loaded on only one rotated component. A Q-matrix based on the components analysis was then constructed. The simulated Q-matrix is shown in Table 3. The resulting simple structure from the components analysis of the simulated responses is shown in Table 4. It can be seen that based on this Q-matrix, there are 5 skill combinations hence 5 components. More specifically, component 1 corresponds to skill 1 because only items measuring skill 1 load on it. The same logic applies to the other components: component 2 represents skill 2; component 3 represents skills 1and 2; component 4 corresponds to skills 2 and 3; and finally, component 5 corresponds to skill 3. The results indicate that the method fully recovered the original Q-matrix used to simulate the data. This was to be expected because the data were simulated and the Q-matrix was known.

Table 3: The Q-matrix (Simulation)

K1	K2	K3
1	0	0
1	0	0
1	0	0
1	0	0
1	0	0
0	1	0
0	1	0
0	1	0
0	1	0
1	1	0
1	1	0
1	1	0
1	1	0
1	0	0
1	0	0
1	0	0
1	0	0
0	1	1
0	1	1
0	1	1
0	1	1

The original Q-matrix for the fraction subtraction data from de la Torre and Douglas (2004) is shown in Table 5. The items measure eight attributes: 1) Convert a whole number to a fraction, 2) Separate a whole number from a fraction, 3) Simplify before subtracting, 4) Find a common denominator, 5) Borrow from a whole number, 6) Column borrow to subtract the second numerator from the first, 7) Subtract numerators, and 8) Reduce answers to the simplest form. Using a sample of n=136 for the components analysis, the simple structure obtained is shown in Table 6. Because with real data there is skills overlap, each of the items loading on a component was examined to determine the skills they measure. Using these skills, a Q-matrix is then constructed. The components analysis resulted in 6 skills for the fraction subtraction data as shown in Table 7. These skills were 1) borrowing, 2) subtract numerators, 3) subtract whole numbers, 4) find a common denominator, 5) subtract a fraction from itself, and 6) put fraction into proper form. The next steps are to perform a confirmatory analysis to test how accurately the constructed Q-matrix classifies examinees into mastery/non-mastery classes. Preliminary findings indicate improved classification accuracy.

Table 4: Components Solution of the Simulated Data

	Component				
	1	2	3	4	5
item1	**0.851**	-0.032	0.026	0.018	0.061
item2	**0.83**	0.01	-0.009	-0.01	0.013
item3	**0.855**	0.008	0.016	-0.014	-0.026
item4	**0.871**	0.023	-0.049	0	-0.035
item5	**0.848**	0.001	0.033	-0.006	0.005
item6	0.006	**0.881**	-0.017	0.038	0.03
item7	-0.006	**0.927**	0.044	-0.081	-0.057
item8	0.002	**0.893**	-0.018	0.028	0.037
item9	0.014	**0.881**	-0.026	0.03	0.034
item10	0.025	-0.009	-0.027	0.015	**0.846**
item11	0.041	0.033	0.021	0.002	**0.794**
item12	0.017	0.038	0.016	0.003	**0.811**
item13	-0.026	0.009	0.006	-0.016	**0.862**
item14	0.114	-0.038	**0.762**	0.085	0.022
item15	0.016	-0.032	**0.797**	0.051	0.021
item16	-0.029	0.018	**0.872**	-0.042	-0.029
item17	-0.024	0.028	**0.842**	-0.074	0.005
item18	0.006	0.089	0.029	**0.745**	-0.006
item19	0.012	-0.128	-0.101	**0.911**	0.032
item20	-0.009	0.02	0.013	**0.788**	-0.022
item21	-0.031	0.114	0.086	**0.726**	-0.018

Table 5: Fraction Data Q-matrix from de la Torre and Douglas (2004)

		Skills							
		K1	K2	K3	K4	K5	K6	K7	K8
Item	1	0	0	0	1	0	1	1	0
Item	2	0	0	0	1	0	0	1	0
Item	3	0	0	0	1	0	0	1	0
Item	4	0	1	1	0	1	0	1	0
Item	5	0	1	0	1	0	0	1	1
Item	6	0	0	0	0	0	0	1	0
Item	7	1	1	0	0	0	0	1	0
Item	8	0	0	0	0	0	0	1	0
Item	9	0	1	0	0	0	0	0	0
Item	10	0	1	0	0	1	0	1	1
Item	11	0	1	0	0	1	0	1	0
Item	12	0	0	0	0	0	0	1	1
Item	13	0	1	0	1	1	0	1	0
Item	14	0	1	0	0	0	0	1	0
Item	15	1	0	0	0	0	0	1	0
Item	16	0	1	0	0	0	0	1	0
Item	17	0	1	0	0	1	0	1	0
Item	18	0	1	0	0	1	1	1	0
Item	19	1	1	1	0	1	0	1	0
Item	20	0	1	1	0	1	0	1	0

Table 6: Components Solution for the Real Data

	The Rotated Matrix												
	Component												
	1	2	3	4	5	6	7	8	9	10	11	12	13
V1	0.21	**0.841**	0.071	0.064	0.087	-0.176	-0.035	-0.037	-0.053	-0.046	-0.115	0.211	-0.073
V2	-0.045	**0.831**	-0.045	-0.082	-0.012	0.165	-0.017	0.103	-0.034	0.06	0.065	-0.027	0.077
V3	-0.184	**0.959**	0.028	-0.099	-0.077	0.016	0.005	-0.04	0.049	0.109	0.09	0.003	0.083
V4	0.102	0.055	0.026	0.036	0.102	-0.024	0.039	-0.072	0.066	-0.053	**0.899**	-0.096	0.014
V5	0.025	0.018	0.012	0.012	0.046	-0.015	0.023	**0.967**	0.014	0.007	-0.073	-0.022	0.015
V6	-0.002	0.145	0.003	0.064	0.01	0.07	0.056	-0.02	0.193	-0.144	-0.085	**0.85**	-0.009
V7	0.052	0.15	**0.648**	0.193	-0.021	0.073	-0.011	0.022	-0.193	0.175	0.071	-0.052	-0.177
V8	-0.053	-0.02	-0.01	-0.019	**0.99**	0.009	-0.022	0.045	0.006	-0.011	0.096	0.01	-0.011
V9	-0.081	-0.108	0.009	**0.986**	-0.018	0.014	-0.021	0.011	0.035	0.048	0.033	0.066	0.088
V10	**0.846**	0.05	0.091	0.034	-0.081	0.049	-0.004	0.055	0.158	-0.08	0.039	-0.26	0.005
V11	**0.94**	-0.019	-0.15	-0.03	-0.108	-0.052	-0.028	0.078	-0.091	-0.074	0.219	0.229	-0.059
V12	-0.019	-0.034	0.028	-0.023	-0.023	-0.014	**0.981**	0.024	-0.008	0.02	0.041	0.062	-0.012
V13	0.092	0.102	-0.007	0.093	-0.012	-0.049	-0.012	0.016	-0.017	-0.083	0.015	-0.011	**0.91**
V14	0.04	0.008	-0.006	0.014	0.008	**0.988**	-0.013	-0.013	0.036	-0.098	-0.024	0.069	-0.046
V15	0.034	0.082	**1.007**	-0.008	0.03	-0.044	0.064	-0.03	0.109	-0.231	-0.067	-0.065	0.018
V16	0.053	-0.015	-0.003	0.035	0.007	0.038	-0.008	0.013	**0.788**	0.217	0.067	0.215	-0.018
V17	0.111	0.112	-0.066	0.047	-0.01	-0.088	0.016	0.006	0.186	**1.003**	-0.046	-0.142	-0.076
V18	**0.574**	0.022	-0.034	0.027	0.108	0.115	0.118	-0.091	-0.168	0.272	-0.117	-0.009	0.148
V19	-0.029	-0.189	**0.693**	-0.144	-0.061	0.002	-0.068	0.059	-0.027	0.244	0.11	0.201	0.124
V20	**0.802**	-0.121	0.161	-0.119	0.076	0.011	-0.047	-0.062	0.095	0.135	-0.096	0.014	0.057

Note: Extraction Method: Principal Component Analysis. Rotation Method: Promax with Kaiser Normalization.

Table 7: Reconstructed Q-matrix for the Fraction Subtraction Data

		Skills					
		K1	K2	K3	K4	K5	K6
Item	1	0	1	0	1	0	0
Item	2	0	1	0	1	0	0
Item	3	0	1	0	1	0	0
Item	4	1	0	0	0	1	0
Item	5	0	1	1	1	0	1
Item	6	0	1	0	0	0	0
Item	7	1	1	0	0	0	0
Item	8	0	0	0	0	1	0
Item	9	0	1	1	0	0	0
Item	10	1	1	1	0	0	0
Item	11	1	1	1	0	0	0
Item	12	0	1	0	0	0	1
Item	13	1	1	0	1	1	0
Item	14	0	1	1	0	1	0
Item	15	1	1	0	0	0	0
Item	16	0	1	1	0	0	0
Item	17	1	1	1	0	0	0
Item	18	1	1	1	0	0	0
Item	19	1	1	0	0	0	0
Item	20	1	1	1	0	0	0

4. Discussion

The components analysis method for Q-matrix development appears to be a viable and useful step in generating a Q-matrix when skill sets are measured by more than one item. Once items have been developed by content specialists, these items should be pilot tested and a task analysis using components analysis conducted to finalize the Q-matrix before items are used operationally. A caveat is that items must be designed to be diagnostic of specific skills covering narrow content areas with a sufficient number of items measuring each skill set. In such circumstances, the components analysis method can be a powerful tool for augmenting theory in the development of the Q-matrix.

The component representation of items fitting the DINA model and the DINA model representation itself are based on different definitions of a "dimension." A dimension in the components representation corresponds to a set of skills that are necessary and sufficient for a high probability of passing the item. A dimension in the DINA model corresponds to a single skill that is necessary, but not necessarily sufficient by itself, for a high probability of passing the item. This mismatch between the definitions of a dimension means no one-to-one match between components and DINA dimensions. Still, if there are multiple items corresponding to each (or most) skill sets, a components solution, in combination with theory and content expert knowledge, can be a useful exploratory tool for deriving a Q-matrix.

In summary, the process of deriving a Q-matrix described in this study begins by conducting a components analysis on binary examinee responses. The items loading on each of the components are examined to identify blocks of items with the same skill set. If the components solution is sensible in terms of the cohesiveness of the skill sets measured by the items that load on each component, the Q-matrix is then constructed. The constructed Q-matrix has a dimension for each skill identified in the components solution. A confirmatory analysis is then conducted to test the performance of the constructed Q-matrix in examinee classification accuracy. Because CDA is heavily reliant on the correct specification of the Q-matrix (DeCarlo, 2011), the exploratory technique investigated in this study has important educational implications. The proposed technique complements theory in Q-matrix development to ensure that the Q-matrix is a valid representation of the skills that a test intends to measure, hence valid skills scores. These scores, in turn, can help improve classroom instruction because teachers can individualize instruction on a student-by-student basis. Improved classroom instruction that is learner-centered should lead to higher academic achievement, which is the ultimate goal of education.

5. References

DeCarlo, L. T. (2011). On the analysis of fraction subtraction data: The DINA model, classification, latent class sizes, and the Q-matrix. *Applied Psychological Measurement, 35,* 8-26.

de la Torre, J. (2008). An empirically based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement, 45*(4), 343.

de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika, 69*(3), 333.

de la Torre, J., & Douglas, J. A. (2008). Model evaluation and multiple strategies in cognitive diagnosis: An analysis of fraction subtraction data. *Psychometrika, 73*(4), 595.

Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement, 26,* 333–352.

Hartz, S. M. (2002). A Bayesian framework for the Unified Model for assessing cognitive abilities: Blending theory with practicality. Unpublished doctoral dissertation, University of Illinois, Champaign, IL.

Henson, R. A., & Templin, J. L. (2007, April). *Large-scale language assessment using cognitive diagnosis models.* Paper presented at the annual meeting on the National Council on Measurement in Education, Chicago IL.

Henson, R., Templin, J. L., & Douglas, J. (2007). Using efficient model based sum-scores for conducting skills diagnoses. *Journal of Educational Measurement, 44*(4), 361.

Leighton, J. P., & Gierl, M. J. (2007). Review of cognitively diagnostic assessment and a summary of psychometric models. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive Diagnostic Assessment for Education: Theory and Applications* (pp. 3–18). New York: Cambridge University Press.

Liu, Y., Douglas, J. A., & Henson, R. A. (2009). Testing person fit in cognitive diagnosis. *Applied Psychological Measurement, 33*(8), 579.

Rupp, A. A., & Templin, J. L. (2008). The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Educational and Psychological Measurement, 68*(1), 78.

Tatsuoka, K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement, 20,* 345–354.

Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods, 11*(3), 287.