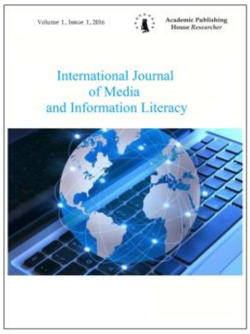
Published in the Slovak Republic
International Journal of Media and Information Literacy
Has been issued since 2016.
E-ISSN: 2500-106X
2018, 3(2): 53-65

DOI: 10.13187/ijmil.2018.2.53
www.ejournal46.com



Assessing Network Media Literacy in China: the Development and Validation of a Comprehensive Assessment Instrument

C.K. Cheung [a, *], Yin Wu [b]

[a] University of Hong Kong, China
[b] Zhejiang University of Media and Communications, China

Abstract

Network media literacy is the foundation of Internet usage and builds sustainable development that can help people to participate more easily in knowledge societies. Nevertheless, no validated and standardised test assesses the level of network media literacy. Therefore, this study established and calibrated an instrument for use in network media literacy research and practice. Items were formed based on a composite conceptual model and administered to the general population across most of the country. The psychometric properties of the questionnaire were examined using multidimensional item response theory. Differential item functioning was used to exclude the items with distorted ability estimates. Almost all of the remaining items showed good discrimination and difficulty parameters based on the fitted model with three stable dimensions. This study created a thorough questionnaire called the general network media literacy test (GNMLT), with scoring determined in relation to classical test theory. The GNMLT is a valid and reliable measure for assessing the network media literacy of Chinese individuals. Practitioners could use the scale before implementing literacy promotion and education.

Keywords: Network media literacy, multidimensional item response theory, differential item functioning, score, China.

1. Introduction

Today, we are witnessing a major shift in information and communication technology, as the Internet is becoming one of the most dominant media. Contrasted with the traditional, linear, hierarchical, logical, rule-governed conventions of print and audiovisual media, the Internet and mobile networks are characterised by multimedia texts, hypertextuality, anarchic organisations, synchronous communication, interactivity, cultural diversity and inclusivity (Livingstone, 2004). As network devices increasingly augment our brains and senses, our knowledge is becoming more widely distributed, and we are becoming 'the sum of our connections and relationship' (Pegrum, 2014). People are increasingly coming to live in a network society structured around network or digital communications. Thus, the Internet has emerged as the 'dominant cultural logic' of our time. Correspondingly, a new form of literacy is emerging, which studies have termed 'computer literacy' or 'Internet literacy' (Livingstone, 2004).

Media literacy is traditionally defined as the ability to understand, analyse, evaluate and create media messages in a wide variety of forms (Aufderheide, 1993), or similarly referred to as the ability to access, analyse, evaluate and communicate messages in a wide variety of forms

* Corresponding author
E-mail addresses: cheungck@hku.hk (C.K. Cheung)

(Young, 2015). Moreover, the plurality of literacy, or the idea that different kinds of literacy are related to the acquisition and application of literacy in particular social contexts, has come to be recognised (UNESCO, 2004) as well-worn terms like 'visual literacy', 'digital literacy' and 'information literacy' have more recently been joined by 'multiliteracies', 'attention literacy' and even 'network literacy'. It is becoming increasingly evident that navigating overlapping personal, social and professional networks – all linked together technologically by the Internet – requires a level of network literacy (Pegrum, 2014). To empower people to make effective use of networks, capitalising on their benefits while avoiding some of their more obvious pitfalls, it is essential to begin fostering media literacy that focuses on the network (Doyle et al., 2012).

The rationale for a media literacy test. Apart from the traditional domains of media literacy, network media literacy introduces several key points that bear consideration. Critical thinking is a particular construct for Internet literacy, as several international studies have supported the link between media literacy and critical thinking. The Feuerstein group examined media literacy as a means to develop critical thinking in children, and concluded that as pupils increased their experience with their media literacy programmes, they showed greater gains proportionally in media analysis and critical thinking skills (Feuerstein, 1999). In addition, Silverblatt, Miller, Smith and Brown (Silverblatt et al., 2014) identified the primary element of media literacy as 'a critical thinking skill that enables audiences to develop independent judgments about media content'. Media literacy is first and foremost about applying critical thinking skills when facing high-capacity network information. Moreover, network spaces like the Internet create forms of literacy that go against traditional understandings of what constitutes content or an interaction; thus, critical literacy becomes emancipatory (Gounari, 2009).

Attitude towards media is an important factor in media literacy competency. Attitude reflects one's desire to positively influence an individual's motivations and perceptions (Powell et al., 2011). Numerous studies have suggested that media has a significant effect on an individual's attitude (Chen et al., 2013). In a networked age, digital literacy is depicted as the awareness, attitude and ability of individuals to appropriately use digital tools (Martin, Grudziecki, 2006), which highlights the intrinsic characteristic of attitudes towards communication, expression and social action (Goodfellow, 2011). It is clear that new media like the Internet have had a dramatic effect on society by modernising peoples' traditional values and attitudes. Thus, conceptualisations of attitude that involve network media literacy align with current thinking about what generates positive individual outcomes.

A comprehensive measure of network media literacy. UNESCO has deconstructed the media and information literacy (MIL) competency standard into three aspects: (i) access and retrieval, (ii) understanding and evaluation and (iii) creation and sharing. In addition, MIL competency is a combination of three cognitive elements: attitudes (rights, principles, values and attitudes), knowledge and skills. These combined cognitive elements are more relevant in a complex environment.

Following UNESCO's deconstruction, network media literacy could be defined as a set of competencies that empowers an individual to access, retrieve, understand, evaluate and use to create and share information and media content via networks in a critical and effective way. Based on this framework and the factors listed in the text, our ultimate conceptual model consisted of three main domains: media skills, media critical thinking and media attitude.

A questionnaire is needed to define and operationalise an individual's network media literacy. However, a validated and standardised test of this literacy is not available. Therefore, in this study, we aimed to develop and explore the psychometric properties of a questionnaire on network media literacy by applying a multidimensional analysis to validate its use. Additionally, we investigated the most optimal methodology for calculating the scores.

2. Materials and methods

Item generation. To generate an item pool, we searched the literature for relevant instruments measuring media, information or digital literacy. We also consulted members of the China Media Literacy Society and experts in the area of media literacy for additional items not represented in the existing measures.

To ensure face validity, 10 students each from the primary school, middle school and university levels were asked to comment on the questionnaire items and give feedback on areas such as the formulation and relevance of the questions and appropriateness of the responses. The items were adapted based on this information. This resulted in a pool of 71 items, after duplicate items were deleted. Both the students and expert panel confirmed the face validity of the scale.

Participants. The participants in the study were gathered from a generally diverse population from 30 provinces in China. Their willingness to participate was ascertained through the processes dictated by the Institutional Review Board at Zhejiang Media Literacy Institute.

Demographic questions sought information about gender, age, occupation and level of education. The next section presented a randomised series of 71 statements (GNMLT). The questions were assessed on the same 5-point Likert scale from 1 (Strongly Disagreed) to 5 (Strongly Agreed).

We classified the participants into two groups based on residence: large and medium-sized cities (municipalities, provincial capitals and prefectures) and small cities (counties, villages and towns). In addition, according to the social development statuses of the different regions based on the Human Development Index (HDI), we classified the participants into two groups: HDI+ and HDI-, which comprised participants with an HDI level above and below the average level of the country (HDI=0.693), respectively. Age grouping was based on the traditional adult age (18).

3. Discussion

We calculated descriptive statistics for the participants' demographics, including frequencies and proportions, to provide preliminary statistical information. We then calibrated the scale based on item response theory (IRT) (Edelen, Reeve, 2007).

A critical assumption of IRT is unidimensionality, which we tested in two ways. First, we used modified parallel analyses (MPA) incorporated into the ltm package to test for the probability of unidimensionality ($\alpha = 0.05$). Second, we used the rule of thumb that the ratio of the first to the second 'eigenvalue' should be above three (Ismail et al., 2013).

Next, we sought to examine the psychometric properties of the GNMLT using multidimensional item response theory (MIRT) in the R software package (Chalmers, 2012), which returned α and β parameters for each item per dimension. Typically, the α parameter indicates the discriminative power of that item. Items with higher scores are better able to discriminate between literate and illiterate individuals. The β parameter presents the difficulty of the imminent dimensions. Higher absolute β scores indicate easier items, while scores towards zero indicate more difficult items. Squared β parameters indicate the degree to which a certain item explains the variance within a certain dimension (Ismail et al., 2013).

We used differential item functioning (DIF) to investigate the degree to which some of the items gave advantages or disadvantages to certain participant groups in relation to the estimates of their ability. The rationale of DIF analyses is to identify items that distort the ability estimates for participants and thus jeopardise the correctness of overall test measurements (Magis et al., 2010). Items that are identified to distort test measurement are referred to as having DIF.

We determined the optimal number of dimensions for the test by comparing the different models with varying numbers of dimensions performed based on MIRT. First, we conducted a deviance test (chi-square test). Second, we compared the differences in Akaike's information criterion (AIC) and the Bayesian information criterion (BIC). Typically, lower AIC and BIC values indicate a better fitting model. When selecting the appropriate number of dimensions, statistical solutions and content-driven arguments must be weighted (Ismail et al., 2013).

Scoring. Finally, to make the GNMLT ready for practical use, we calculated test scores using the classical test theory (CTT) approach and compared them with test scores generated via IRT analysis. In CTT, test scores are simply a sum of the number of correctly answered items (each correctly answered item is assigned one point). In contrast, IRT scores account for the level of difficulty per item. Next, we compared the scores generated via the CTT approach and those generated via IRT analysis using a Pearson correlation test. If the CTT scores were a close approximation of the IRT-derived test scores, then they were considered potentially favourable for practical proposes, as they were easier to calculate. We identified the cut-off point of the scores and evaluated its sensitivity and specificity.

4. Results

Participant characteristics. Figure 1 shows the socio-demographic characteristics of the participants. The project recovered 6,478 samples, and the effective recovery rate was 91.96 %, including 55.3 % for males and 43.8 % for females. The mean age was 28±11.4yrs. The survey involved 30 provinces, with 30.9 % of the participants coming from Zhejiang, 6.3 % from Shanxi, 5.7 % from Shandong. Urban and rural areas accounted for 66.9 % and 33.1 % of the participants, respectively. In terms of education, the primary level and below accounted for 6.8 % of the participants, the middle-school level (including vocational high school and technical school) accounted for 33.2 % and the university level and above accounted for 40 %. Cross-analysis revealed no significant difference in educational composition between the men and women. The sample was fairly representative of the general population in China.

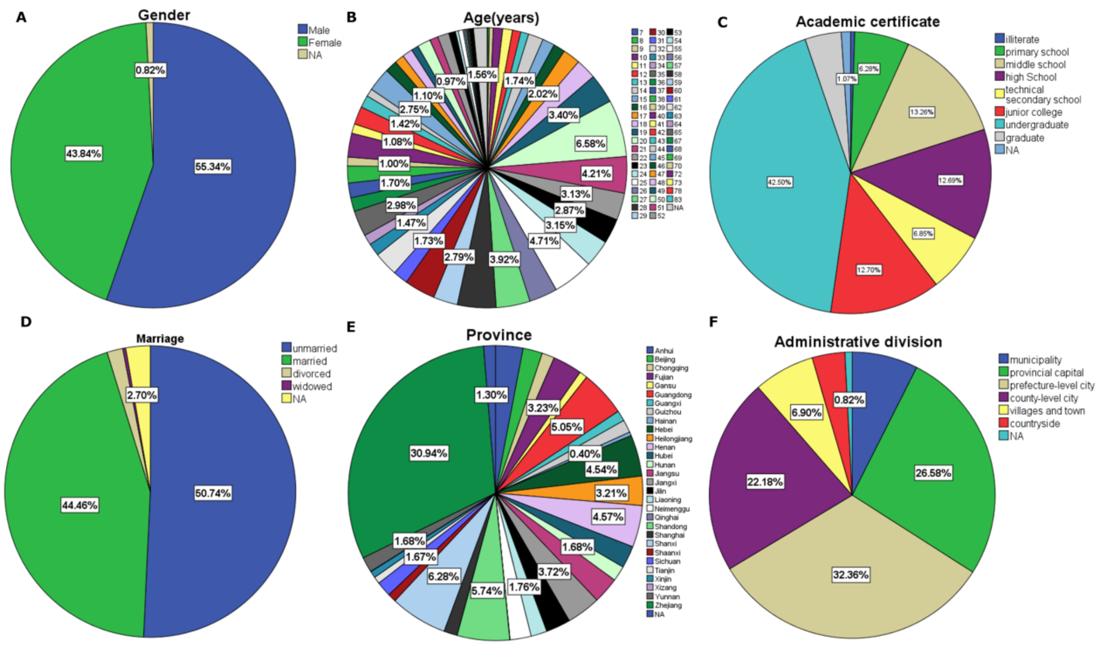


Fig. 1. Participant characteristics

Profile of the participants in this study. 55.3 % were male and 44 .8% were female. The largest proportion (41.8 %) was aged 26–30 years, followed by those aged 21–25 years (39.6 %). 87.9 % of the participants were unmarried. 70.7 % of the participants were college- or university-educated, and 17.2 % had received a graduate school education.

Differential item functioning. With a given latent trait, estimates of item characteristics should hold true regardless of the group being tested. The importance of literacy competence underlines the strong need to understand the gender gap in literacy achievement (Schwabe et al., 2015). Socio-demographic variables such as subject residence are held to affect network users' willingness and ability to productively use network media. This effect can create a participation divide between distinct region groups (Hoffmann et al., 2015).

We began by using DIF to develop a broader applicable scale. We identified 48 items as candidates for deletion from the list and subsequent analysis due to distorted ability estimates, as indicated by a significant DIF (Figure 2). In addition, three items were deleted between region groups. For example, participants from the countryside (rural) received relatively lower ability estimates on item 1 ('Acquire information knowledge through a network') than the reference groups (P = 0.1939). This distortion also held for items 2 ('Freedom to show themselves on the Internet') (P = 0.1204) and 4 ('Satisfy curiosity through a network') (P = 0.0732). Therefore, these three items were excluded from further analyses.

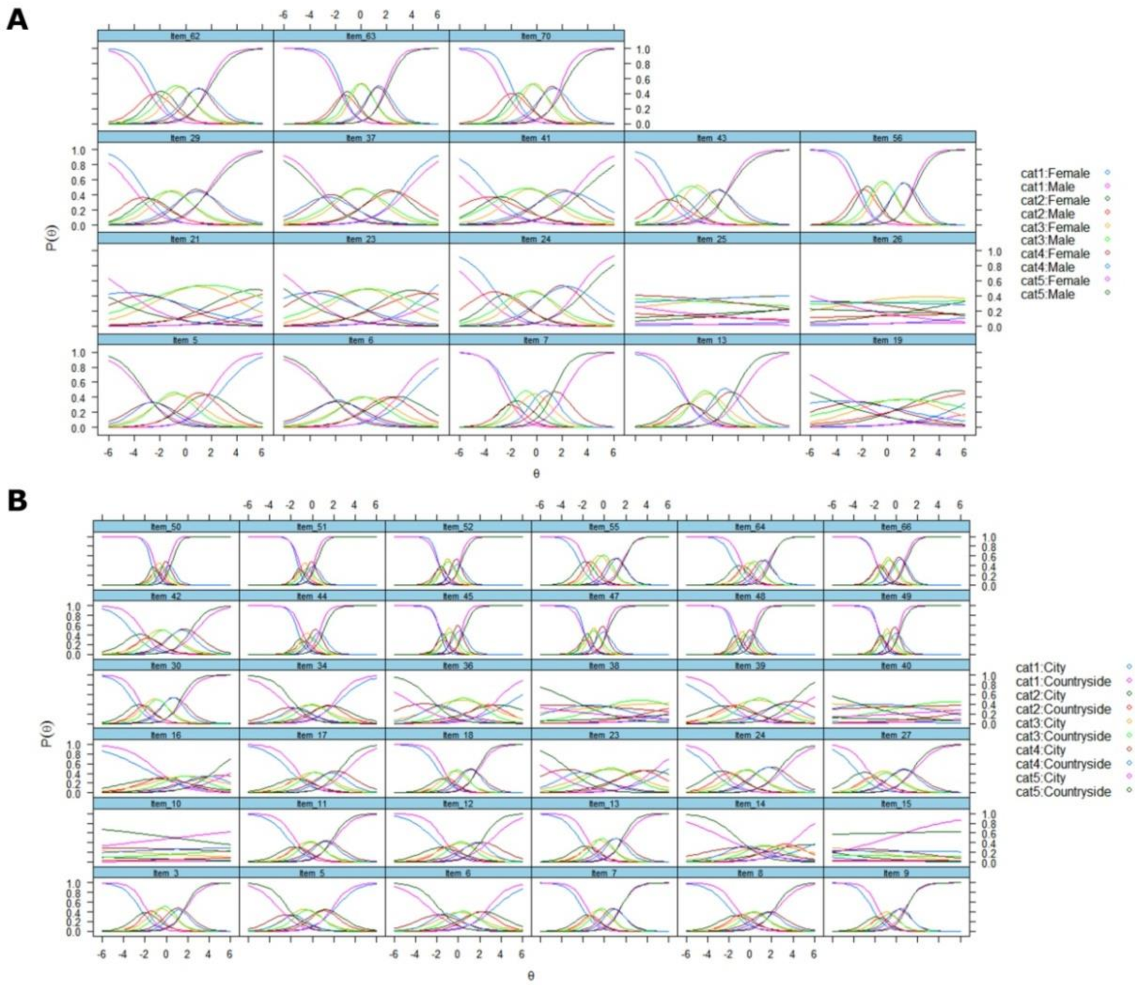


Fig. 2. Differential item functioning in gender grouping (A) and region grouping (B)

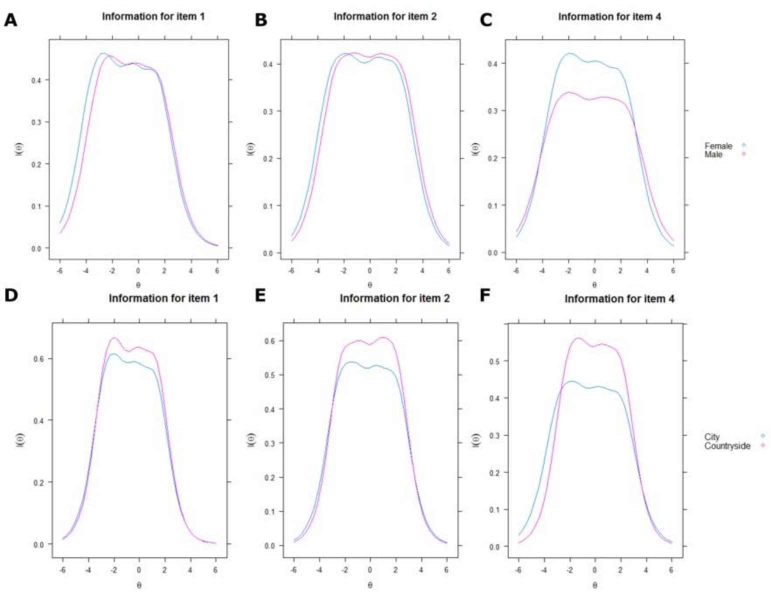


Fig. 3. Information trace for Items 1, 2 and 4. θ_p is the ability of a person.
I(θ) refers to the corresponding test division

Multidimensional item response theory. Before fitting an appropriate IRT model, the critical assumption of unidimensionality was assessed. The test for unidimensionality using MPA on all of the participants together was significant (P = 0.009). The ratio of the first to second eigenvalue was 7.342/2.191 = 3.35.

Given these findings in support of multidimensionality, MIRT was applied to the remaining 20 items. A one-dimensional model postulating general literacy was tested against a two- to six-dimensional model. Table 1 shows that the difference between the models up to model 5 is significant at the $\alpha = 0.05$ level.

Table 1. Comparing multidimensional item response theory models

Model	Log-likelihood	AIC	BIC	Comparing models
Model 1	-144443.9	289087.8	289757.5	
Model 2	-139902.4	280042.9	280839.9	(Model 1 versus Model 2; χ^2 = 9,082.893, d.f. = 19, P < 0.001)
Model 3	-137543.2	275360.4	276277.9	(Model 2 versus Model 3; χ^2 = 4,718.467, d.f. = 18, P < 0.001)
Model 4	-137154.1	274616.1	275647.5	(Model 3 versus Model 4; χ^2 = 778.263, d.f. = 17, P < 0.001)
Model 5	-136859.5	274059.1	275197.6	(Model 4 versus Model 5; χ^2 = 589.093, d.f. = 16, P < 0.001)
Model 6	-136934.3	274238.6	275477.6	(Model 5 versus Model 6; χ^2 = -149.502, d.f. = 15, P = 1.000)

Table 1 shows the log-likelihood, Akaike's information criterion (AIC) and Bayesian information criterion (BIC) parameters for the fitted models with the MIRT() function. The final column shows the comparisons between the nested models using a deviance test (chi-square statistic, degrees of freedom, P-value). The models reflect the number of dimensions tested ('Model 1' contains one dimension, 'Model 2' contains two dimensions, etc.).

Table 2. Factor summary of Model 5

	Factor1	Factor 2	Factor 3	Factor 4	Factor 5	Communalities
Item 20	.058	-.009	.159	-.383	.051	0.221
Item 22	.126	.101	.122	-.523	.008	0.489
Item 28	.035	.300	.034	-.276	-.142	0.342
Item 31	-.045	.864	.027	-.009	-.006	0.737
Item 32	.023	.935	-.010	.022	.013	0.861
Item 33	.086	.543	-.022	-.026	.015	0.348
Item 35	.044	.232	.021	-.181	-.099	0.185
Item 46	.359	.009	.242	.013	.064	0.279
Item 53	.618	.059	-.054	-.255	-.057	0.610
Item 54	.770	.014	.005	-.099	.008	0.672
Item 57	.685	.008	.127	.016	.126	0.524
Item 58	.798	.033	.019	.048	-.053	0.699
Item 59	.683	.028	.028	.060	-.142	0.598
Item 60	.041	.010	.098	-.014	-.526	0.348
Item 61	.400	.014	.097	.052	-.400	0.547
Item 65	.197	-.038	.538	.117	.086	0.393
Item 67	-.008	.019	.808	-.038	-.016	0.680

Item 68	.004	.062	.745	-.045	-.030	0.630
Item 69	-.029	-.076	.753	.030	-.005	0.500
Item 71	.017	.059	.692	-.041	-.035	0.563
SS loadings	2.912	2.091	2.682	.634	.525	

The five-factor model is summarised in Table 2, which displays the factor loading for the items in each dimension. From a conceptual viewpoint, the fourth and fifth factors contain items that cover varying subject areas.

To cross-validate the structure, we performed a principal component analysis (PCA) using oblique rotation to test the factor structure. The contribution rates of Factors 4 and 5 were 4.77 % and 4.14 %, respectively, and the eigenvalues were less than 1. The first three factors accounted for 57.70 % of the total variance.

Overall, the three-dimensional model had a better fit and more coherent content. In addition, the factor communality estimate of item 35 in Factor 2 was relatively low (0.185), indicating the small homogeneity of this item. Hence, items 20 ('Network information can lead the trend'), 22 ('Networks can influence behaviour'), 35 ('I want others to be honest with me on a network'), 60 ('I can distinguish between harmful information') and 61 ('I can measure the media information') were excluded from further analyses.

Table 3. Item characteristics on the three subscales of the 15-item test

Items	Short item description	α1 (SE)	α2 (SE)	α3 (SE)	β*1* (SE)	β*2* (SE)	β*3 (SE)*	β*4 (SE)*
Media attitude								
Item 28	Control of network information is necessary	0.504 (0.022)	0.834 (0.027)	0.169 (0.036)	2.126 (0.108)	4.001 (0.113)	4.496 (0.118)	3.893 (0.119)
Item 31	On-line advertising exaggerated the effect of goods	0.547 (0.05)	2.781 (0.096)	0.365 (0.099)	5.443 (0.188)	9.214 (0.266)	9.749 (0.287)	6.362 (0.256)
Item 32	The information content on the network is not always correct	0.963 (0.081)	4.175 (0.163)	0.478 (0.167)	8.078 (0.326)	13.48 (0.504)	14.443 (0.55)	9.775 (0.448)
Item 33	Sometimes there is a bias on the network	0.384 (0.023)	1.155 (0.037)	0.164 (0.042)	2.497 (0.094)	4.381 (0.101)	4.302 (0.108)	2.083 (0.116)
Media critical thinking								
Item 46	Information classification	0.718 (0.024)	0.234 (0.020)	0.702 (0.024)	1.463 (0.074)	2.777 (0.079)	2.702 (0.084)	1.643 (0.085)
Item 53	Evaluation of the different views	1.723 (0.039)	0.884 (0.045)	0.471 (0.067)	3.414 (0.132)	6.242 (0.165)	7.027 (0.18)	4.754 (0.173)
Item 54	Network delivery implications	2.200 (0.047)	0.782 (0.052)	0.861 (0.076)	4.878 (0.161)	8.243 (0.202)	8.379 (0.214)	4.762 (0.203)
Item 57	Know the popular term on the network	1.38 (0.033)	0.404 (0.032)	0.825 (0.044)	2.736 (0.105)	4.952 (0.12)	4.988 (0.127)	2.9 (0.129)
Item 58	Distinguish between harmful information	2.204 (0.048)	0.645 (0.055)	0.912 (0.074)	4.482 (0.154)	7.691 (0.195)	7.935 (0.208)	4.838 (0.196)
Item 59	Measure the media	1.683 (0.037)	0.479 (0.043)	0.699 (0.057)	3.608 (0.139)	6.236 (0.161)	6.373 (0.169)	3.94 (0.163)
Media skill								
Item 65	Production of network video	0.555 (0.028)	0.052 (0.021)	1.175 (0.021)	1.326 (0.061)	2.053 (0.068)	1.177 (0.074)	1.247 (0.1)

Item 67	Focus on hot events and comments	0.873 (0.053)	0.462 (0.028)	2.169 (0.031)	3.605 (0.117)	5.576 (0.143)	4.832 (0.147)	1.528 (0.148)
Item 68	Share articles	0.831 (0.047)	0.5 (0.027)	1.826 (0.03)	3.132 (0.118)	5.249 (0.142)	5.061 (0.147)	2.548 (0.143)
Item 69	Follow celebrity and forward	0.418 (0.048)	0.125 (0.03)	1.862 (NA)	2.227 (0.081)	3.174 (0.097)	2.200 (0.101)	0.691 (0.123)
Item 71	Share information across multiple media	0.68 (0.045)	0.319 (NA)	1.886 (NA)	2.979 (0.109)	4.887 (0.129)	4.248 (0.134)	1.369 (0.135)

Items are paraphrased for brevity. This table displays slopes transformed into a varimax-rotated factor loadings metric: item discrimination parameter [α (SE)] and item difficulty parameters for the three respective dimensions [β (SE)].

After the aforementioned analyses, we constructed a three-dimensional GNMLT test based on 15 items. Table 3 presents the three-dimensional solution with the factor labels. The factor labels 'media attitude' (MA), 'media critical thinking' (MC) and 'media skill' (MS) reflect the content of the respective factors.

Bifactor validation. Next, we performed a confirmatory factor analysis (CFA) of the factor structures. Figure 4 shows the path diagram for the model.

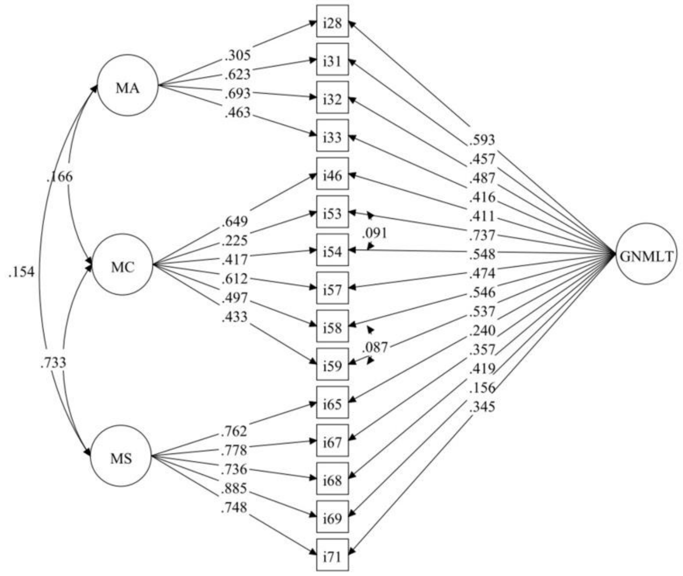


Fig. 4. Standardised path diagram

Results of bifactor CFA of the GNMLT. MA = media attitude; MC = media critical thinking; MS = media skill.

Considering the fit indices of the model, we calculated the ratio of χ2/df as 2.14. In addition, Table 4 presents the other fit indices and evaluates them in line with the related literature. This indicated a good fit with the proposed model (Muthén, Asparouhov, 2012). Thus, we confirmed that the model has three factors.

Table 4. Evaluation of fit indices under CFA

Indice	Sample statistic	Perfect fit	Good fit	Decision	Rationale
χ2/df	2.14	χ2/df≤ 2	χ2/df≤ 3	Good fit	(Sideridis et al., 2014)
RMSEA (Root Mean Square Error Of Approximation)	0.037	RMSEA≤ 0.05	RMSEA≤ .08	Perfect fit	(Iacobucci, 2010)
RMR (Root mean square residual)	0.050	RMR≤ 0.05	RMR≤ 0.08	Perfect fit	(Brown et al., 2006) (Hu, Bentler, 1999)
SRMR	0.022	SRMR≤0.05	SRMR≤ 0.08	Good fit	(Brown et al., 2006)

(Standardised root mean square residual)					(Hu, Bentler, 1999)
CFI (Comparative Fit Index)	0.983	CFI≥ 0.95	CFI≥ 0.90	Perfect fit	(Hu, Bentler, 1999) (Sivo et al., 2006)
TLI (Tucker-Lewis index)	0.974	TLI≥0.95	TLI ≥ 0.90	Perfect fit	(Iacobucci, 2010)

When performing reliability analysis as a result of the CFA, we calculated the Cronbach's alpha (α) internal consistency coefficient for the total scale made up of three factors as α=0.908, that for the factor of media attitude as α=0.824, that for the factor of media critical thinking as α=0.877 and that for the factor of media skill as α=0.878.

Test scores. Test scores represent the aggregate of the item responses. To validate the stability of the GNMLT test, we grouped the samples by age, province, gender and region. The distribution of the test scores showed a close-fitting score curve in all subgroups (Figure 4 A-D).

Level of education directly affects an individual's literacy (Hobbs, 1998; Kellner, Share, 2005). Figure 4E shows the total distribution of literacy, with the samples grouped by education level. The total scores for the extreme groups are 38.41 ± 9.63 for the illiterate group and 56.58 ± 9.76 for the postgraduate group, with Cohen's d = 1.87 (Jacobson, Truax, 1991).

We conducted receiver operating characteristic analysis to determine the possible cut-off of the GNMLT. We used education level above university to identify the participants with adequate literacy. The curve for the GNMLT showed that scores ≥ 50 on the scale had a sensitivity of 70.31 % and a specificity of 60.18 % for predicting adequate literacy (Figure 4F).

CTT has been replaced by IRT (Wirth, Edwards, 2007), and its scoring is relatively simple. In this study, the IRT and CTT scores were highly correlated (r = 0.897, P = 0.000). Figure 4G shows the CTT score curve. The cut-off value was estimated at 8 points. The area under the receiver operating characteristic curve for predicting adequate health literacy is 0.684 (95 % CI, 0.671–0.695, p <0.001). The GNMLT curve shows that scores ≥ 8 on the CHLCC had a sensitivity of 58.26 % and a specificity of 70.62 % for predicting adequate literacy.

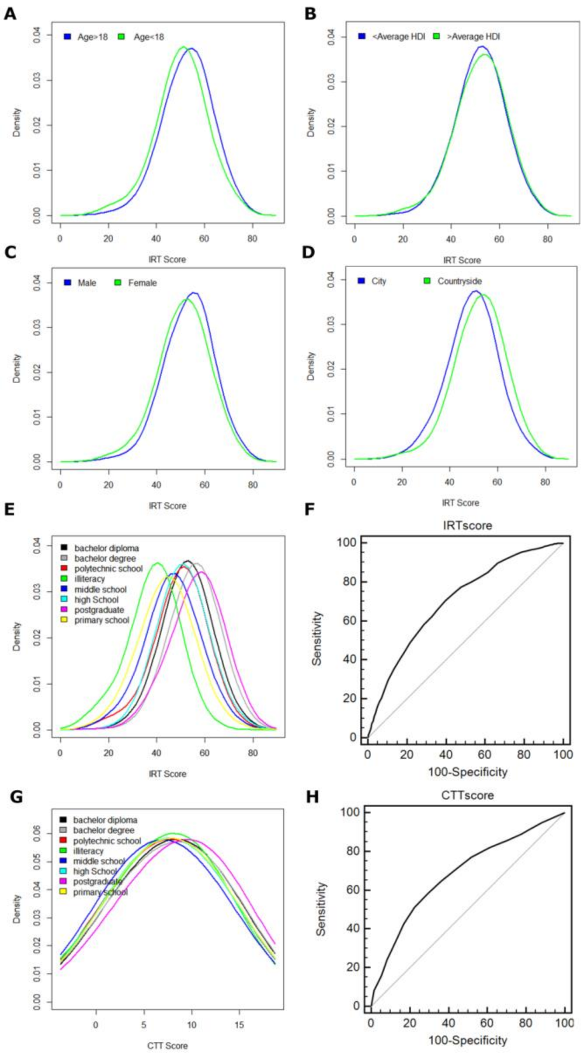


Fig. 5. Test scores and cut-off values

The X-axis is the test score, and the Y-axis reflects the fraction of the subject's score. (A-D) Test score distribution in each subgroup according to age, HDI, gender and region, respectively. (E) IRT test score distribution based on education level. (F) ROC curve for IRT test score. (G) CTT test score distribution based on education level. (H) ROC curve for CTT test score.

5. Conclusion

Over the past decade or so, the Internet and mobile network technology have transformed multiple facets of life in society that have changed our work and leisure patterns. Indeed, network media literacy is becoming indispensable.

However, although the field of media literacy is growing in terms of both interest and participation, relatively little quantitative research has examined media literacy evaluation (Bergsma, Carney, 2008; Cheung, 2016). An important reason for this is the challenge of measuring media literacy (Arke, Primack, 2009). As Scharrer states, "The results of participation in media literacy curricula are not often explicitly defined and measured, but there is a generalized notion about what these outcomes are" (Scharrer, 2002: 354). To show the value of literacy, tools must be developed and possessed to accurately measure and report different literacy results.

Of note, media literacy is an umbrella concept. In the understanding, manufacture and coordination of culture represented by symbol, text, geometry, sound and image, media literacy is emerging and spreading in the digital information signal (Peek, Beresin, 2016). In view of this, this initial measurement tool was based on current media literacy research and focused on the more modern form of Internet communication. Based on the MIL Assessment Framework of UNSEO, attitudes, knowledge, skills and critical thinking represent logical starting points in the development of network media literacy measures.

As the latent characteristics of individual literacy are indirect, this study focused on multidimensional IRT analysis. Most of the items showed good discrimination parameters based on the model that was fitted, indicating that individuals with various literacy levels could be discriminated adequately.

After serial modelling and verification by PCA and bifactor CFA, the final scale had three solid dimensions. Hence, we concluded that a three-dimensional test was the best solution in this case. We loaded three subscales together in the model and mapped them onto literacy: media attitude, media critical thinking and media skill. These subscales exhibited satisfactory reliability. Taken together, the items reflected the overall construct of network media literacy in the Chinese population.

This scale also measured network media literacy in general. For this purpose, larger and more diverse samples were used in this study, which covered most regions of China. The calibration process included differential item functioning and factor analysis, resulting in a final 15 questions. Moreover, the scale was stable when validating the different subgroups (Fig. 4A-D) and valuable in targeting the general population.

The final version of the GNLMT consisted of 15 questions that required about 5 minutes to complete, showing the increase in time needed to accomplish the scale. Short scales like this should encourage the measurement of literacy before any educational intervention. The GNLMT could be used to assess levels of literacy in population surveys.

We determined the optimal scoring method for the practical use of the test on the population. Determining test scores for the 15 items using both CTT and IRT showed very high correlations. In light of this, the CTT method is preferable, because the calculations are easily performed by hand: each correct response receives a score of 1 and each incorrect response a score of 0.

This study has some limitations. First, it used convenience sampling, which might have generated a selection bias. However, given the filtration of items with DIF and the valid and cohesive three-dimensional structure, the sample distribution and distinctiveness of the different subgroups did not indicate any flaws in the sample. Second, some of the information used in the scale was context- or language-specific. To use the GNLMT in other regions, researchers may need to re-examine some items to ensure that they are suitable for these contexts. It would also be worthwhile to study the use of literacy scales across different ethnic groups. Third, the GNLMT did not cover the entire media literacy perspective, and was not meant to supplant traditional media literacy skills. Individuals must have traditional literacy skills other than network literacy to expand their knowledge (Young, 2015). However, as network media literacy research develops alongside media literacy studies in the mobile network world, additional measurement tools and areas will need to be developed to keep measurement efforts current and applicable.

In conclusion, the present study outlines the development and validation of the 15-item GNLMT. This newly developed instrument provides a user-friendly measure of network media literacy for use in the general Chinese population.

References

Arke, Primack, 2009 – *Arke, E.T., Primack, B.A.* (2009). Quantifying media literacy: Development, reliability, and validity of a new measure. *EMI Educ Media Int, 46*(1), 53-65.

Aufderheide, 1993 – *Aufderheide, P.* (1993). Media literacy: A report of the national leadership conference on media literacy.

Bergsma, Carney, 2008 – *Bergsma, L.J., Carney, M.E.* (2008). Effectiveness of health-promoting media literacy education: A systematic review. *Health Educ Res, 23*(3), 522-542.

Brown et al., 2006 – *Brown, D.J., Cober, R.T., Kane, K., Levy, P.E., Shalhoop, J.* (2006). Proactive personality and the successful job search: A field investigation with college graduates. *Journal of Applied Psychology, 91*(3), 717-726.

Chalmers, 2012 – *Chalmers, R.P.* (2012). MIRT: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(6), 29.

Chen et al., 2013 – *Chen, Y.-C.Y., Kaestle, C.E., Estabrooks, P., Zoellner, J.* (2013). US children's acquisition of tobacco media literacy skills: A focus group analysis. *Journal of Children and Media, 7*(4), 409-427.

Cheung, 2016 – *Cheung, C.K.* (2016). The Future of Media Literacy Education in China: The Way Forward. In Cheung C.K. (Ed.) *Media Literacy Education in China*, Springer, 173-179.

Doyle et al., 2012 – *Doyle, F., Watson, R., Morgan, K., McBride, O.* (2012). A hierarchy of distress and invariant item ordering in the General Health Questionnaire-12. *J Affect Disord, 139*(1), 85-88.

Edelen, Reeve, 2007 – *Edelen, M.O., Reeve, B.B.* (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Qual Life Res, 16-1*, 5-18.

Feuerstein, 1999 – *Feuerstein, M.* (1999). Media literacy in support of critical thinking. *Journal of Educational Media, 24*(1), 43-54.

Goodfellow, 2011 – *Goodfellow, R.* (2011). Literacy, literacies and the digital in higher education. *Teaching in Higher Education, 16*(1), 131-144.

Gounari, 2009 – *Gounari, P.* (2009). Rethinking critical literacy in the new information age. *Critical Inquiry in Language Studies, 6*(3), 148-175.

Hobbs, 1998 – *Hobbs, R.* (1998). The seven great debates in the media literacy movement. *Journal of Communication, 48*(1), 16-32.

Hoffmann et al., 2015 – *Hoffmann, C.P., Lutz, C., Meckel, M.* (2015). Content creation on the Internet: A social cognitive perspective on the participation divide. *Information, Communication & Society, 18*(6), 696-716.

Hu, Bentler, 1999 – *Hu, L.T., Bentler, P.M.* (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling-a Multidisciplinary Journal, 6*(1), 1-55.

Iacobucci, 2010 – *Iacobucci, D.* (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology, 20*(1), 90-98.

Ismail et al., 2013 – *Ismail, S.Y., Timmerman, L., Timman, R., Luchtenburg, A.E., Smak Gregoor, P.J., Nette, R.W., et al.* (2013). A psychometric analysis of the Rotterdam Renal Replacement Knowledge-Test (R3K-T) using item response theory. *Transpl Int, 26*(12), 1164-1172.

Jacobson, Truax, 1991 – *Jacobson, N. S., & Truax, P.* (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol, 59*(1), 12-19.

Kellner, Share, 2005 – *Kellner, D., & Share, J.* (2005). Toward critical media literacy: Core concepts, debates, organizations, and policy. *Discourse: Studies in the Cultural Politics of Education, 26*(3), 369-386.

Livingstone, 2004 – *Livingstone, S.* (2004). Media literacy and the challenge of new information and communication technologies. *The Communication Review, 7*(1), 3-14.

Magis et al., 2010 – *Magis, D., Beland, S., Tuerlinckx, F., De Boeck, P.* (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behav Res Methods, 42*(3), 847-862.

Martin, Grudziecki, 2006 – *Martin, A., Grudziecki, J.* (2006). DigEuLit: Concepts and tools for digital literacy development. *Innovation in Teaching and Learning in Information and Computer Sciences, 5*(4), 1-19.

Peek, Beresin, 2016 – *Peek, H.S., Beresin, E.* (2016). Reality check: How reality television can affect youth and how a media literacy curriculum can help. *Acad Psychiatry, 40*(1), 177-181.

Pegrum, 2014 – *Pegrum, M.* (2014). *Mobile learning: Languages, literacies and cultures*. Springer.

Powell et al., 2011 – *Powell, R.B., Stern, M.J., Krohn, B.D., Ardoin, N.* (2011). Development and validation of scales to measure environmental responsibility, character development, and attitudes toward school. *Environmental Education Research, 17*(1), 91-111.

Scharrer, 2002 – *Scharrer, E.* (2002). Making a case for media literacy in the curriculum: Outcomes and assessment. *Journal of Adolescent & Adult Literacy, 46*(4), 354-358.

Schwabe et al., 2015 – *Schwabe, F., McElvany, N., Trendtel, M.* (2015). The school age gender gap in reading achievement: Examining the influences of item format and intrinsic reading motivation. *Reading Research Quarterly, 50*(2), 219-232.

Sideridis et al., 2014 – *Sideridis, G., Simos, P., Papanicolaou, A., & Fletcher,* J. (2014). Using structural equation modeling to assess functional connectivity in the brain: Power and sample size considerations. *Educational and Psychological Measurement, 74*(5), 733-758.

Silverblatt et al., 2014 – *Silverblatt, A., Miller, D.C., Smith, J., Brown, N.* (2014). Media literacy: Keys to interpreting media messages. Santa Barbara, CA: ABC-CLIO.

Sivo et al., 2006 – *Sivo, S.A., Fan, X.T., Witta, E.L., Willse, J.T.* (2006). The search for "optimal" cutoff properties: Fit index criteria in structural equation modeling. *Journal of Experimental Education, 74*(3), 267-288.

Sector, 2004 – *Sector, U.E.* (2004). The plurality of literacy and its implications for policies and programs: Position paper [J]. Paris: United National Educational, Scientific and Cultural Organization, 2004, 13.

Wirth, Edwards, 2007 – *Wirth, R.J., Edwards, M.C.* (2007). Item factor analysis: Current approaches and future directions. *Psychol Methods, 12*(1), 58-79.

Young, 2015 – *Young, J.A.* (2015). Assessing new media literacies in social work education: The development and validation of a comprehensive assessment instrument. *Journal of Technology in Human Services, 33*(1), 72-86.